THE UNIVERSITY *of* EDINBURGH

Investigation of Protein-Ion Interactions by Mass Spectrometry and Ion Mobility Mass Spectrometry

Yana Berezovskaya

PhD

The University of Edinburgh

2012

To the memory of my Mother

"Confidence is what you have before you understand the problem."

Woody Allen

Acknowledgements

I would like to thank my supervisor Dr Perdita E. Barran for making my distant-future-when-more-important-things-are-taken-care-of plan to pursue a PhD a reality. I am also grateful for a generous conference allowance, both domestic and international. Thank you for preferentially tasking me with handling most delicate and expensive (!) parts of the instruments from day one of my PhD project.

I am grateful to my collaborator Prof. Dek Woolfson and his people (Craig Armstrong and Aimee Boyle) at the University of Bristol, who provided synthetic peptides for the studies described in Chapters 3 and 4. Thank you for all things related to in-solution analysis and your valuable feedback and discussions.

I would like to thank the members of the Barran research group, past and present, for their help, support and cheerful company.

I am grateful to all the technical personnel, both internal and external, for their vast research-enabling capabilities. A lot of time has been saved and frustration spared thanks to your knowledge and professionalism.

I am indebted to my parents for the wonderful childhood they gave me, for developing my personality, and for their numerous sacrifices to ensure my education.

My thanks and eternal gratitude are directed to my husband, Dr Lev Sarkisov for his most professional guidance, discussions, support, advice and reality checks. Thank you for invaluable Linux scripts that made the analysis of the data in Chapter 5 possible. Thank you for always being there for me.

I would like to thank my cat Loki for being an unceasing source of positive emotions. Thank you for getting me out of my chair at the desk during my thesis write-up: your persistence and accurate timing in demanding food and attention have kept my days structured.

Declaration

This thesis is submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Edinburgh. Unless otherwise stated, this work is my own and has not been submitted for any other degree or professional qualification.

……..

Yana Berezovskaya

2012

Abstract

Protein-ion interactions play an important role in biological systems. A considerable number of elements (estimated 25 – 30) are essential in higher life forms such as animals and humans, where they are integral part of enzymes involved in plethora of cellular processes. It is difficult to overestimate the importance of thorough understanding of how protein-ion interplay affects living cell in order to be able to address therapeutic challenges facing humanity. Presented to the reader's attention is a gas-phase biophysical analysis of peptides' and proteins' interactions with biologically relevant ions (Zn^{2+} and I^-). This investigation provides an insight into conformational changes of peptides and proteins triggered by ions.

Mass spectrometry and ion mobility mass spectrometry are used in this work to probe peptide and protein affinities for a range of ions, along with conformational changes that take place as a result of binding. Observation of peptide and protein behaviour in the gas phase can inform the investigator about their behaviour in solution prior to ionisation and transfer from the former into the latter phase. Wherever relevant, the gas-phase studies are complemented by molecular dynamics simulations and the results are compared to solution phase findings (spectroscopy).

Two case studies of protein-ion interactions are presented in this thesis. Firstly, sequence-to-structure relationships in proteins are considered *via*

protein design approach using two synthetic peptide-based systems. The first system is a synthetic consensus zinc finger sequence (vCP1) that is responsive to zinc: it adopts a zinc finger fold in the presence of Zn^{2+} by coordinating the metal ion by two cysteines and two histidines. This peptide has been selected as a reference for the zinc-bound state and a simple model to refine the characterisation method in preparation for analysis of a more sophisticated second system – dual conformational switch. This second system (ZiCop) is designed to adopt either of the two conformations in response to a stimulus: zinc finger or coiled coil. The reversible switch between the two conformational states is controlled by the binding of zinc ion to the peptide. Interactions of both peptide systems with a number of other divalent metal cations (Co^{2+}, Ca^{2+} and Cu^{2+}) are considered also, and the differences in binding and switching behaviour are discussed. Secondly, protein-salt interactions are investigated using three proteins (lysozyme, cytochrome *c* and BPTI) using variable temperature ion mobility mass spectrometry. Ion mobility measurements were carried out on these proteins with helium as the buffer gas at three different drift cell temperatures – 'ambient' (300 K), 'cold' (260 K) and 'hot' (360 K), and their conformational preferences in response to HI binding and temperature are discussed.

The work of this thesis has been covered in the following publications and presentations.

Papers:

1 Berezovskaya, Y.; Armstrong, C. T.; Boyle, A. L.; Porrini, M.; Woolfson, D. N.; Barran, P. E. (2011). Metal binding to a zinc-finger peptide: a

comparison between solution and the gas phase. *Chemical Communications* 47, 412-414 (work described in Chapter 3).

2 Berezovskaya, Y.; Porrini, M.; Nortcliffe, C.; Armstrong, C. T.; Woolfson, D. N.; Barran, P. E. (2012). Ion mobility mass spectrometry as a tool for synthetic biology: a case study on the zinc finger fold versus coiled coil interactions. *In preparation* (work described in Chapter 4).

3 Berezovskaya, Y.; Porrini, M.; Barran, P. E. (2012). The effect of salt adductation on the conformations of three model proteins is revealed by variable temperature ion mobility mass spectrometry. *Submitted to International Journal of Mass Spectrometry* (work described in Chapter 5).

Presentations:

4 BMSS (British Mass Spectrometry Society) Annual Meeting, Cardiff, UK, 2011, oral: "Metal ions and mass spectrometry assist synthetic biology".

5 ARF (Analytical Research Forum), Manchester, UK, 2011, oral: "Ion-mobility mass spectrometry: a tool for synthetic biology".

6 ASMS (American Society for Mass Spectrometry) Annual Meeting, Denver, USA, 2011, oral: "Metal ions and mass spectrometry assist synthetic biology".

7 University of Edinburgh School of Chemistry 2nd year Postgraduate Student Presentation, Edinburgh, UK, 2010, oral: "Ion-mobility mass

spectrometry: a tool for synthetic biology" (second place in competition) .

8 ACS (American Chemical Society) Fall Meeting, Boston, USA, 2010, oral: "Metal binding of model Cys2His2 zinc-finger peptides".

9 BMSS (British Mass Spectrometry Society) Annual Meeting, Cardiff, UK, 2010, poster: "Metal binding of model Cys2His2 zinc finger peptide in the gas phase".

10 ARF (Analytical Research Forum), Loughborough, UK, 2010, poster: "Metal binding of model Cys2His2 zinc finger peptide in the gas phase".

11 IMSC (International Mass Spectrometry Conference), Bremen, Germany, 2009, oral: "Peptides designed to switch their conformation by metal binding".

12 ARF (Analytical Research Forum), Kent, UK, 2009, poster: "Ion-mobility mass spectrometry study of a peptide designed to switch its conformation by metal binding".

Abbreviations

3D	three-dimensional
Ac	acetate (salt)
APCI	atmospheric pressure chemical ionisation
API	atmospheric pressure ionisation
ATD	arrival time distribution
AUC	analytical ultracentrifugation
BPTI	bovine pancreatic tryptic inhibitor
CCS	collision cross section
CD	circular dichroism
CID	collision-induced dissociation
CS	charge state
CV	compensation voltage; collision voltage
DC	direct current
DMS	differential mobility spectrometry
DNA	deoxyribonucleic acid
EDD	electron detachment dissociation
EHSS	exact hard sphere scattering
ESI	electrospray ionisation
ETD	electron transfer dissociation
FAB	fast atom bombardment
FAIMS	high-field asymmetric waveform ion mobility spectrometry

Fmoc	fluorenylmethyloxycarbonyl (chloride): protecting group for amines
HBTU	O-benzotriazole-N,N,N′,N′-tetramethyl-uronium (hexafluoro-phosphate): peptide coupling reagent
HPLC	high-performance liquid chromatography
i.d.	internal diameter
IE	injection energy
IM	ion mobility
IM-MS	ion mobility mass spectrometry
IMS	ion mobility spectrometry
IPA	isopropyl alcohol
LD	laser desorption; linear drift
LDT	linear drift tube
MALDI	matrix-assisted laser desorption/ionisation
MCP	microchannel plate
MD	molecular dynamics
MoQ-ToF	mobility Q-ToF
MS	mass spectrometry
MS/MS	tandem mass spectrometry
nESI	nano-electrospray ionisation
NMR	nuclear magnetic resonance
o.d.	outer diameter
PA	projection approximation
PDB	protein database
Pp	partner peptide
Q	quadrupole

RF	radio frequency
RMSD	root mean square deviation
RNA	ribonucleic acid
SA	simulated annealing
SLD	soft laser desorption
SRIG	stacked ring ion guide
TCEP	tris(2-carboxyethyl)phosphine: reducing agent for disulphide bonds
TDC	time-to-digital converter
TIC	total ion chromatogram
TM	trajectory method
TMV	tobacco mosaic virus
ToF	time of flight
TWIMS	travelling wave ion mobility spectrometry
UV	ultraviolet
vCP1	variant consensus peptide 1
XIC	extracted ion chromatogram
XRC	X-ray crystallography
ZiCop	zinc finger coiled coil peptide

Table of contents

Chapter 1

Introduction

1.1 Peptides, proteins and their interactions with ions

A large proportion of the living cell is comprised of proteins: they are responsible for over a half (~60%) of the 'dry' mass of the cell[1]. The role of proteins in living organisms is all-encompassing: they provide structure, facilitate transport of small molecules and ions, catalyse chemical reactions, convert energy, mediate the immune response and assist in perpetuation (and adaptation) of the cell's genetic code[2]. In order to function correctly, it is necessary for proteins to fold into unique native conformations. For a polypeptide chain, to reach the native folded state by a random search of all possible conformations would be prohibitively long, whereas most proteins arrive to their native state within a second. A number of different theories emerged to explain how proteins fold[3], partly due to advances in experimental techniques that enable monitoring of folding events. Among key factors in protein folding are individual amino acid residues

contributions that 'direct' the polypeptide chain into favourable pathways. These amino acids interact with each other and such 'external' factors as ions, small molecules and other proteins (*e.g.* chaperones), channelling the folding process into a series of energetically favourable intermediate states, resulting in protein folding on biologically-relevant time-scales. There is a number of different types of interactions involved in protein folding, including hydrogen bonds, electrostatic interactions, salt bridges, van der Waals forces, hydrophobic interactions and steric effects[4]. When studying proteins, it is often very difficult to decouple the effects of one type of interaction from another, as they are usually found in a complex interplay.

The main focus of this thesis will be on electrostatic interactions, even though they inevitably coexist with other interactions. Many biological macromolecules, including proteins, rely on electrostatic interactions for their function. Charged residues on a protein surface play an important role in attracting ionic ligands from the surrounding solvent[5; 6; 7]. The net charge of a protein, as well as the distribution of that charge (*i.e.* negatively and positively charged side chains) on the molecule's surface, is an essential factor in determining the strength of the protein-ion interactions[8]. The work presented in this thesis employs mass spectrometry based methods to study electrostatic interactions in proteins.

Two case studies of protein-ion interactions will be presented in this thesis. Firstly, sequence-to-structure relationships in proteins will be considered *via* protein design approach. Here, a conformational transition will be observed in response to zinc cation binding to a model synthetic peptide system. Secondly, effect of iodide anion from the Hofmeister series on three model proteins will be studied.

In the next section (***1.1.1***), a brief overview of protein structural features and folding behaviour will be given. This will be followed by a review of zinc ion-mediated conformational switching in proteins considered through protein design approach (section ***1.1.2***). Then, the effect of salts on proteins will be considered in the framework of anions from Hofmeister series (section ***1.1.3***). Finally, the relevance of gas-phase method to address the aforementioned scientific challenges will be discussed (section ***1.2*** and ***1.3***) and compared to number of complementary techniques (section ***1.4***).

1.1.1 Protein structure and folding

Proteins are biological polymers of α-amino acid monomeric units. An amino acid has an amino group, a carboxyl group, a hydrogen atom and a side chain group (R) that gives an amino acid its identity, all of which are connected to an α-carbon atom (***Figure 1-1***). The α-carbon atom is a chiral centre rendering amino acids optical activity. All eukaryotic proteins are built from amino acids in L-configuration. Only 20 different side chain groups are universally found in all species populating this planet. The full list of the 20 most common amino acids is given in ***Appendix***, ***Figure A 1***. A wide range of physical properties is offered by the diversity and versatility of chemical composition of these side chains, which ensures the remarkable range of functions mediated by proteins. These physical properties cover size, shape, charge, reactivity and hydrogen bonding capacity.

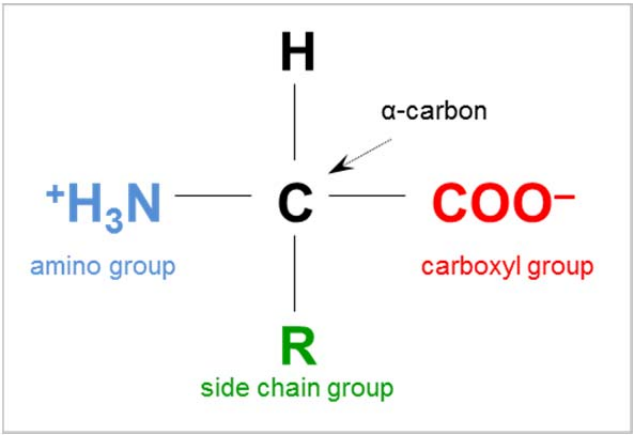


Figure 1-1 *Schematic representation of a general amino acid formula featuring its functional groups: amino (blue), carboxyl (red) and side chain (green). At physiological pH 7.0 both amino and carboxyl groups are ionised yielding a zero net charge.*

Amino acids are enzymatically linked in sequence *via* amino and carboxyl groups in a 'head-to-tail' fashion to form a polypeptide chain. The amino acid sequence constitutes the first level of hierarchy of protein architecture, called primary structure. The polypeptide backbone is arranged in a higher-order configuration, referred to as secondary structure, by interactions between its amino acid residues positioned relatively close to each other along the chain (local folding). These interactions are facilitated by hydrogen bonds and stabilise the structure in a periodic manner giving rise to a number of structural motifs, α-helices and β-sheets being two of the most common ones. Tertiary structure is a spatial arrangement that relies on interactions between the amino acids that are relatively far apart along the polypeptide chain (long-range folding). The interactions that stabilise the tertiary fold are of both covalent (disulphide bridges) and non-covalent nature (hydrophobic and electrostatic interactions). If a protein is constructed from more than one polypeptide chain, each forming a sub-unit of the protein, the interactions between the different polypeptide chains form the quaternary structure (multimeric organisation). Quaternary structure is the

highest level of spatial organisation of a protein, stabilised by various types of covalent and non-covalent interactions.

The specific order of amino acid residues in proteins is in encoded in the cellular DNA. The protein sequence information contained in DNA is copied (transcribed) into messenger RNA (mRNA). Upon leaving the nucleus, mRNA binds to ribosomes where proteins are synthesised by the process of translation as linear chains of typically a few hundred residues long. The correct arrangement at amino acid sequence level is necessary but not sufficient to ensure correct folding of a protein. A complex interplay of physiochemical conditions during and after protein synthesis defines the unique native three-dimensional structure of a protein, ensuring its correct function. The idea of being able to control protein folding has intrigued scientists for over five decades[3; 9; 10; 11; 12], and understanding of sequence-to-structure relationships in proteins have been a cornerstone of efforts across physics, chemistry and biology[13].

1.1.2 Sequence-to-structure relationships studied by protein design: conformational switching triggered by zinc cation binding

Metal ions play an important and diverse role in many biological functions by interacting with proteins; biological phenomena that rely on peptide-metal ion interactions include cellular storage and transport, catalysis and signal transduction. About 40% of the natural amino acids are capable of binding metal ions[14], which renders the presence of such amino acids on a protein surface very probable. Approximately 47% of structurally

determined proteins require metals, with 41% of those containing metals in their catalytic centre[15]. Metals can facilitate protein-protein interactions by binding at the complex interface, stabilising the interaction and orienting the catalytic machinery[16]. Metals are widely involved in protein folding by directing the polypeptide chain to a correct pathway to a functional native conformation by stabilising the resulting structure[17; 18]. Enzyme catalysis is another process where metal ions play a crucial role as they may be part of the a catalytic centre[19; 20]. Perhaps not surprisingly, interactions of metal cations with proteinaceous species have attracted a sustained interest from the protein design community for some time, resulting in plethora of studies pertaining to the characterisation of such systems[21; 22]. And it does not stop here: a substantial amount of work has been carried out based on analysis of existing and generation of new metal-binding sites[23; 24].

The divalent zinc cation is of particular relevance to the presented research. Zinc is the second most abundant transition metal out of the nine that are essential for proper functioning of the human body (preceded by iron)[25] assuming a pivotal role in a variety of biological processes. Examples include its involvement in immune function[26], specific binding of the reproductive system hormone oxytocin to its receptor[27] and transcription factors called zinc fingers[28]. Two chapters (***Chapter 3*** and ***Chapter 4***) of the presented thesis are dedicated to the sequence-to-structure relationship of a designed synthetic peptide study and the role of zinc in directing its fold as well as that of zinc finger proteins.

An important goal of protein design is to replicate, from first principles, active protein folds, and in doing so synthesise novel functional sequences, that may perform useful tasks[29; 30]. Examples include protein-based

biosensors for *in vivo* diagnostics[31] and explosive detection[32]. In addition to practical considerations, protein design is a useful tool to test our understanding of fundamental principles of protein folding and function. Understanding how a protein's primary sequence dictates its fold (conformation) is a key to understanding its function, as the protein's amino acid composition is intrinsically linked to its fold[33; 34]. The main reasons why scientists seek understanding of sequence-to-structure relationships are:

❖ *De novo* design is the ultimate test of our understanding of fundamental principles governing protein folding. The validity and robustness of any theory is its predictive power.

❖ Design of new 'improved' or 'enhanced' characteristics in naturally occurring proteins. Examples of such characteristics may include: improved thermal stability, resistance to wider range of pH, alternative binding selectivity (affinity to different ligands) or specificity (conformational complementarity).

❖ The most challenging task is the design of proteins with novel functions 'from scratch'. For example, design of a protein that has no analogues in the natural world to catalyse chemical reactions.

The usual approach to protein design involves careful consideration of which active fold(s) are required, identification of the minimal components to retain fold, followed by iterative design[35]. The choice of folds to design is highly limited due to the complexity of the task at hand, and therefore is often based on motifs occurring in nature. Only well-characterised motifs with clearly defined folds and short sequences (from 20 to just over 100 amino acid

residues) are amenable to *de novo* design. As a result, only a few protein-folding motifs have been designed successfully *de novo*, and the zinc finger motif is one of them[36; 37].

The following six sub-sections will describe two protein-folding motifs that were studied in this work – zinc finger and coiled coil. The main challenge of the presented *de novo* design was to incorporate a switching functionality between the two folds. The central feature of the design is zinc binding, whereas the role of the coiled coil is to stabilise the zinc finger peptide in the absence of the metal. Similarly to the zinc finger fold, the coiled coil is another structural motif that has been approached successfully by a protein design strategy[38]. The coiled coil motif is the most versatile for *de novo* designed switches as its sequence is highly amenable to being superimposed on many other target sequences[39]. Below, a brief overview of naturally occurring proteins featuring either of these structural motifs is given. This is followed by description of design approaches aimed at obtaining each of these folds. Finally, the particulars of designing duality of sequence and fold are presented.

1.1.2.1 Zinc finger proteins

The zinc finger proteins were the first gene-specific eukaryotic transcription factors to be discovered[40]. They are now a highly studied family of transcription factors and regulatory proteins which are encoded by up to 3% of the human genome[41]. They are found widely in nature where they play an important role in gene regulation and expression and other cellular functions[42]. A zinc finger protein may contain from just 1 to more than 30 repeating zinc finger domains[43; 44] that bind to a range of compounds in the

living cell – proteins, DNA, RNA and small molecules[45]. Robert Roeder and Donald Brown uncovered the protein whilst working on the 5S ribosomal RNA of frog *Xenopus laevis*.[46] It was found that the transcription initiation required the binding of a 40 kDa protein which was referred to as transcription factor IIIA (TFIIIA)[40]. Through deletion mapping it was discovered that TFIIIA interacted with a 50-nucleotide region within the gene called the internal control region. Later Miller began studying TFIIIA[47] and discovered the repeating motif within the structure which was involved with the binding of a DNA and the use of a zinc ion within the structure. This led to the protein fold being referred to as a 'zinc finger'. Misfolding of zinc finger proteins leads to errors in their binding to DNA and function, directly affecting transcription process and resulting in mutation of DNA[37]. These events have been linked to a variety of diseases such as cancer and a series of neurological disorders[48]. A deeper understanding of structural features of the zinc finger domain is imperative to comprehend its function. This includes thorough knowledge of how the particular zinc finger sequence affects its fold, binding strength and specificity of the resulting conformation. Detailed structural and dynamic analysis along with design of artificial zinc finger proteins with new DNA-binding properties and functions could greatly improve our knowledge of the zinc finger biological role as well as potentially aid in construction of a very powerful and useful tool in gene therapy[43].

1.1.2.2 Zinc finger domains

Zinc finger domains exist as a range of spatially distinct structures[44]. Each fold class has a different structural motif although cysteine and histidine residues binding the zinc ion are common features of all. A hydrophobic

cluster is also conserved in most zinc fingers. The zinc ion is a vital component in stabilising the structure of these proteins and in its absence they adopt a different structural conformation – random coil[44]. Depending on the structural features of the zinc-binding site, zinc finger domains can be categorised in eight groups; however there are only three groups to which the majority of all zinc finger folds belong. These groups are summarised in ***Table 1-1***[45], and more details on each of them are given below.

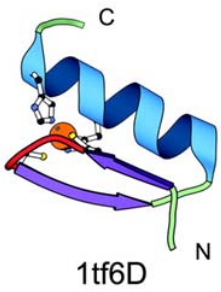
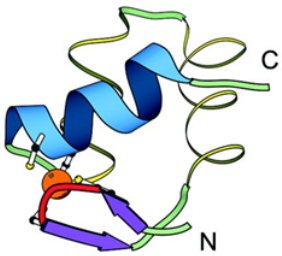
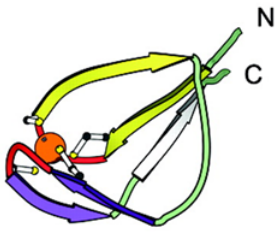
Fold Group	Representative Structure	Ligand Coordination
C_2H_2-like	 1tf6D	Two Zn-binding residues from the β-sheet knuckle and two more from C-terminus of the α-helix
Treble clef	 1i3qJ	Two Zn-binding residues from the β-sheet knuckle and two more from N-terminus of the α-helix
Zinc ribbon	 1iciA	Two Zn-binding residues from each of the two β-sheet knuckles

Table 1-1 *Structural classification of zinc finger motifs featuring three main fold groups. Coordinated zinc ions are shown in orange and metal-binding residues are highlighted. The PDB entry names and chain ID are shown next to each structure. Reprinted (adapted) with permission from Krishna et al.[45], copyright 2003 Oxford University Press.*

The first discovered[47] and most common type of DNA-binding zinc finger motif in nature is the C_2H_2-type domain[28]. This motif is described here in more detail as the focal point of presented work. The C_2H_2-domain usually consists of a sequence of 20-30 amino acids that fold into a motif comprising

two anti-parallel β-sheets on the N-terminus and an α-helix on the C-terminus[43]. The amino acid string of zinc finger motif complies with the consensus sequence (F/Y)-X-C-X_{2-5}-C-X_3-(F/Y)-X_5-Ψ-X_2-H-X_{3-5}-H, where X is any amino acid and Ψ is a hydrophobic amino acid residue[42; 49]. The structure is stabilised by a Zn^{2+} ion[50] that is tetrahedrally coordinated by four amino acid residues – two cysteines (Cys, thiol containing) and two histidines (His, imidazole containing) yielding the ββα-fold (***Table 1-1*** and ***Figure 1-2A***). These two residues, as well as a series of hydrophobic residues, are conserved within most members of this group with Tyr6 (or Phe6), Phe17, and Leu23 being retained[51]. In the resulting structure one of the cysteines is located on the first β-sheet, the second cysteine is on the β-hairpin between the two anti-parallel β-sheets, and the two histidines are localised on the α-helix domain. The α-helical portion of the domain specifically recognises and binds to the major groove of DNA at 3 base pair intervals[45] through several amino acid residues. This leads to the whole zinc finger protein to be wrapped around the DNA strand (***Figure 1-2B***). The most common task associated with this class of zinc finger protein is regulation of transcription[52]. This is achieved by controlling the 'copying' (*i.e.* transcription) of genetic information from DNA to messenger RNA (mRNA). Three base pairs of the DNA (codon) encode one amino acid. The mRNA then carries coding information to ribosomes where proteins are synthesised off the mRNA 'blueprint' (*i.e.* translation).

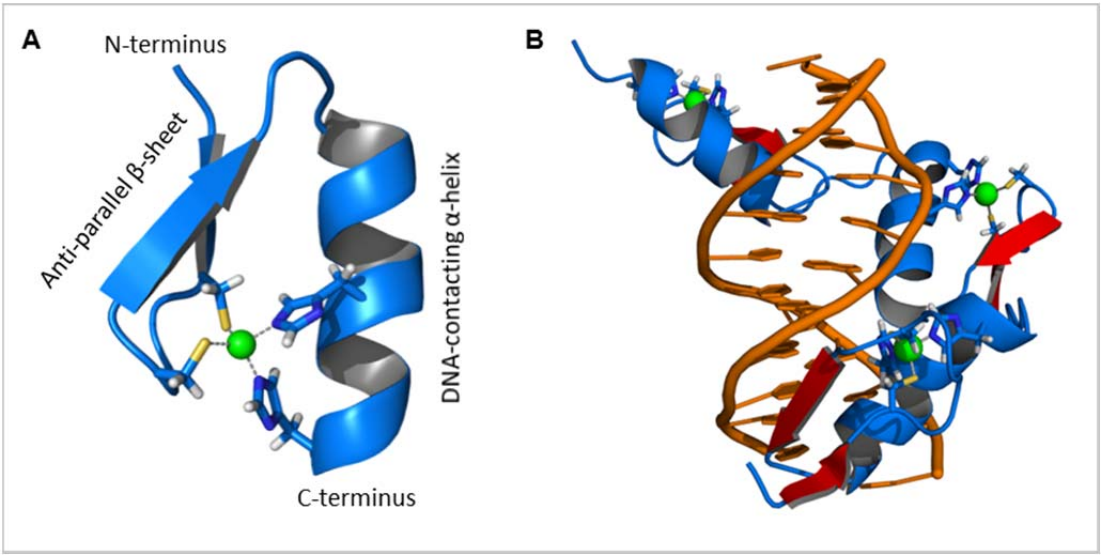


Figure 1-2 *Schematic representation of a C_2H_2-type zinc finger domain featuring the ββα-fold (**A**) and the Zif268 protein containing three zinc finger motifs in complex with DNA (**B**). The protein backbone is shown in blue, zinc ion in green and DNA in orange. Metal-coordinating residues – Cys and His are highlighted. Based on the X-ray structure of PDB 1A1L. Reprinted (adapted) with permission from Thomas Splettstoesser under the terms of the GNU Free Documentation License.*

The treble clef structural motif (***Table 1-1***) contains an α-helix at the C-terminus and a β-hairpin at the N-terminus with each side contributing two amino acid residues for binding of the zinc ion. This class of zinc finger is the most diverse with little sequence or functional conservation between proteins[53]. Nuclear hormone receptors are the best characterised example of this motif[54].

The zinc ribbon fold group (***Table 1-1***) has two zinc knuckles providing the ligands for metal ion binding. The central fold consists of two β-hairpins forming similar binding sub-sites. The N-terminus is commonly referred to as the primary hairpin with the C-terminus as the secondary; the third hairpin often forms hydrogen bonds between these forming an anti-parallel β-sheet. The zinc ribbon is arguably the largest fold group among zinc

fingers with very high sequence diversity and is found in a wide range of proteins[45].

1.1.2.3 Design of zinc finger domains

The Cys_2His_2 zinc finger motif has been the most widely developed basis for genetically engineered proteins targeted to specific sequences of genomic DNA[42]. The majority of these are based upon the transcription factor Zif268 which has three individual sets of zinc finger motifs that bind to a 9 base pair sequence[55]. The structure of Zif268 was solved in 1991 by Pavletich & Pabo[56], initiating a field of research in zinc finger arrays. Most engineered zinc finger proteins have between three and six individual zinc finger motifs and bind from anywhere between nine to eighteen base pairs. It has been established that two to three zinc fingers arranged in tandem are sufficient for the specific DNA binding without additional contribution from participation of any other domains or factors[43].

Zinc-finger motifs have been successfully designed over the last few decades.[36; 37] This success is partially owed to the fact that metal-ion-binding sites are the 'simplest' functional design[57]. Since the first description of a designed tetrahedral Zn(II)-binding site with specified geometry in 1990 by Regan *et al.*[58] there has been sustained interest in such designs[51; 59]. Advances in computer modelling[60] and rapidly growing databases of natural and synthetic zinc finger proteins have made it possible to obtain highly-specific DNA binders by mixing and matching individual modules from various archives. ***Chapter 3*** describes development of gas-phase characterisation platform of the zinc finger peptide derived from the work of Berg group[61], and the findings were published[62].

1.1.2.4 Coiled coil proteins

Coiled coil proteins play an important role in the living cell ranging from transcription to intracellular transport and cellular division[63]. The first characterisation work on coiled coil structure was conducted by two groups simultaneously over 60 years ago – L. Pauling[64] and F. Crick[65]. A coiled coil is a common protein interaction motif wherein 2 – 7 α-helices are wrapped around each other to form a loose left-handed super-helical twist[66]. The naturally most abundant multimers of coiled coils are dimers and trimers. The resulting α-fibril is stabilised by a number of non-covalent interactions: hydrophobic packing, electrostatic and cation-π interactions[66; 67]. The common feature of a coiled coil amino acid sequence is a canonical heptad repeat of alternating hydrophobic (*h*) and polar (*p*) residues, usually arranged in the following order: $(hpphppp)_n$, often denoted as $(abcdefg)_n$. The hydrophobic amino acids are spaced every 3 to 4 residues apart, which allows one hydrophobic residue per turn, since α-helix has approximately 3.6 amino acids per turn. Such arrangement gives rise to an amphipathic structure wherein one side of the α-helix is hydrophobic and the other is polar. ***Figure 1-3*** shows a helical-wheel representation of interactions found in coiled coils. A helical wheel illustrates the properties of alpha helices. The sequence of amino acids is plotted in a rotating manner with the 100° angle of rotation between consecutive amino acids, and the helix propagates along the axis down the plane of the page. The hydrophobic non-polar faces of the α-helices (**a** and **d** residues on the schematic) interact to form a super-coil structure, leaving the polar residues solvent-exposed and potentially available for decoration with functional groups[68]. Of these, **e** and **g** give specificity between the two helices through electrostatic interactions, and the

remaining three positions (**b**, **c** and **f**) must all be hydrophilic, as these will form helical surfaces that are exposed to the solvent.[69; 70]

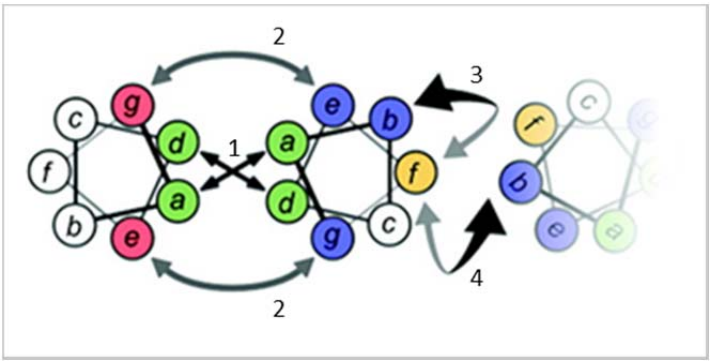


Figure 1-3 *Schematic helical-wheel representation of coiled coil motif featuring typical non-covalent interactions: (**1**) hydrophobic packing; (**2**) charge-charge interactions; (**3**) intra-helical cation-π interactions; (**4**) inter-helical cation-π interactions. Reprinted (adapted) with permission from Gribbon et al.[67], copyright 2008 American Chemical Society.*

The types of interactions taking place in coiled coil motifs are intrinsically non-specific, therefore α-helices tend to be oriented in a variety of ways relative to each other and bundle into fairly misaligned aggregates yielding a wide range of structural architectures (***Figure 1-4***)[71; 72; 73]. Finding ways to control aggregation in a desired manner presents a major challenge to the researchers[63; 74; 75].

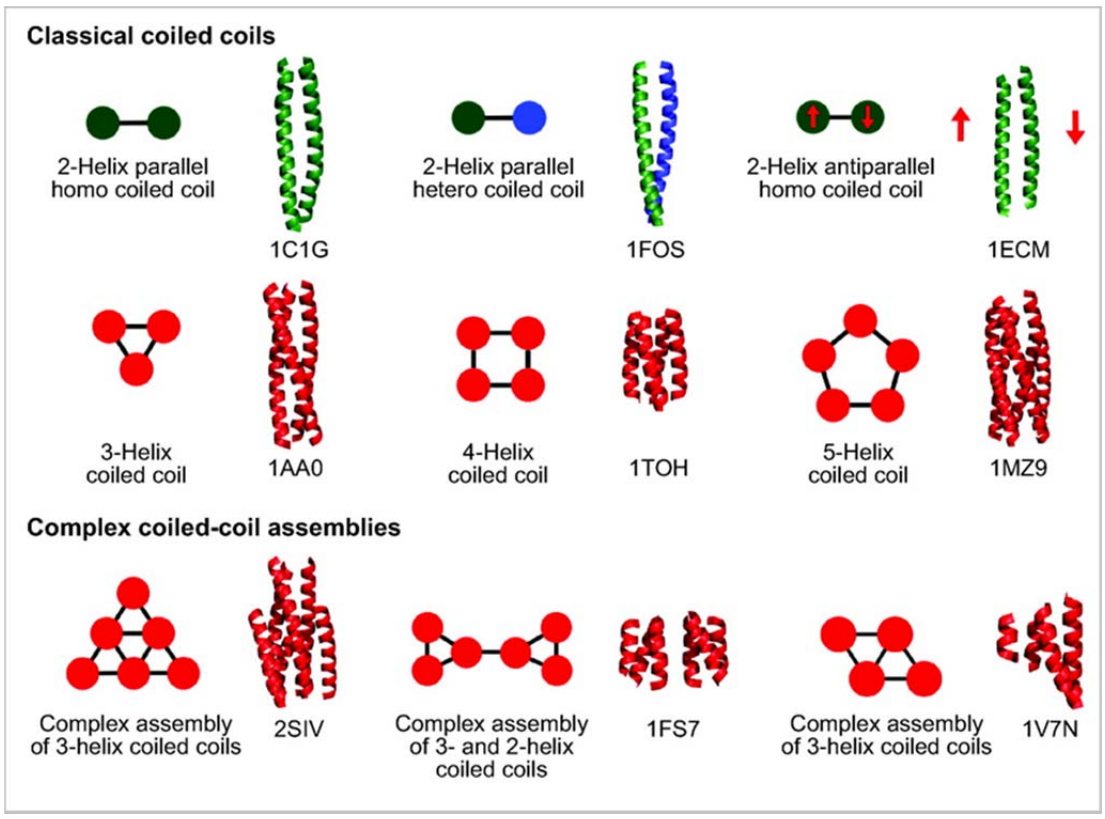


Figure 1-4 *Classical and complex range of coiled coil topologies. Top row – topologies (e.g. parallel, anti-parallel, hetero-typic, or homo-typic); middle row – architectures (multimeric order); bottom row – representative complex coiled coil assemblies (can be represented as different combinations of classical architectures). Reprinted with permission from Bromley et al.[68], copyright 2008 American Chemical Society.*

1.1.2.5 Design of coiled coil domains

Due to its elegant architecture, coiled coil motif is an attractive model for studying protein folding, stability and self-assembly[13; 68; 76]. Along with zinc finger conformations, coiled coils are an inspiration for the *de novo* design of proteins. A promising medical application of biocompatible synthetic coiled coils includes hydrogels for tissue engineering, targeted drug and cell delivery and controlled release[76; 77]. Of particular interest is research into early stages of fibril formation that can be modelled using synthetic coiled coils, and it is difficult to overestimate the impact this knowledge can have

on our understanding of such debilitating human conditions as Alzheimer's and Parkinson's diseases[78].

As has been shown above, coiled coils occur in a range of oligomerisation states and topologies. For this reason they provide excellent components for protein design, where they can be used as potential spacers, general building blocks, and as hubs to co-localise and orient other functional domains[68; 76]. Also, coiled coils have ample room on their external faces potentially available for functionalization or other interactions. The design rules for parallel dimeric coiled coils are well established[79; 80; 81; 82]. Put simplistically, to specify a coiled coil one needs to place hydrophobic amino acids every 3 or 4 residues apart, and fill the rest of the spaces with polar amino acids and the ones with high helical propensity (Ala, Glu, Lys, Gln)[38]. In practice however, the task is more complex, as the differences between multimeric states[80; 83] or topologies of coiled coils[84] are dictated by subtle changes of amino acids comprising them, as well as by the global energy minimum of the constituent subunits. Despite this latter complication, the basic design rules for parallel dimeric coiled coils have been well established based on the vast body of information collected on these sequences[63; 79].

1.1.2.6 Design of switching peptides

There are many faces of naturally occurring conformational switches[85; 86; 87]. Some manifest their switching functionality in transition from a disordered to ordered state, where the transition can either involve the whole protein or a portion of the protein. Others switch between well-defined ordered conformations. Many switches involve a change in oligomerisation state[88; 89]. Designing a novel protein structure with a predictable function is a daunting

task in itself, and when an extra complication of incorporating a switching functionality is added to that, the principles that underline structural ambivalence need to be understood thoroughly.

Building in a conformational switching functionality (usually between two folds) into a single sequence of a designed protein presents an additional challenge. There are two main conditions to be met when designing a switching protein or peptide[39]:

- it has to be stable in two or more well-defined target folded states (each of the individual target structures has to be 'designable');

- it has to respond to a sufficient stimulus to change its conformation (the energy barrier separating the stable states has to be of adequate height to make the conformational transition achievable).

Duality of conformation inevitably must be reflected in duality of sequence: sequence spaces must overlap. This task is usually approached by searching, in addition to sequence space, conformational space in order to find low-energy sequence-structure pairs. This emphasises the importance of structure prediction and refinement in future success in protein design[90]. Another difference between designing a switch and a single state is that, the aim of single-state design is just maximising the stability of the target state, which is often reasonably easy to achieve[91]. In many cases of single-state design, this means that inaccuracies in the energy function can be compensated by placing an emphasis on use of interactions that are known to be favourable. In switch design, this may lead to the conformation to be

trapped in one of the target states, whereas the goal is to achieve a balance between alternative sets of interactions.

An early application of *de novo* approach to combining two distinct folds within a single sequence was reported by Mutter *et al.*[92] This milestone work described a design of amphiphilic peptides capable of conformational transition from α-helix to β-sheet in response to change of pH. A number of different stimuli for the switch have been proposed ranging from temperature, Coulombic interactions, X→N acyl migration (X = O or S) and metal ion interaction[93]. A detailed knowledge of minimal sequence code required for switching protein conformation is essential. Another fine example of such endeavour was reported by Alexander *et al.*[94] who followed a change in protein fold and function mutation by mutation. Their task was to find a mutation that would transform protein's conformation and binding target, and they have shown that the conformational switch occurred *via* a single amino acid substitution with 90% of each side of the switch point populated. Although the switching process described by the authors was not taking place within one sequence ('switching in space'), the findings are very valuable indeed, as they establish sequence-to-fold relationships. A very useful switching polypeptide based on an *in situ* intramolecular O→N acyl group migration was designed by Mimna *et al.*[95] ('switching in time'). The authors investigated whether amyloid formation could be disrupted or reversed through an induced transformation of a β-sheet to an α-helix structure.

Numerous studies investigated specifically zinc finger motifs by designing *de novo* versions[17; 96]. Rational design of metal-binding sites has been only explored in the context of stabilisation of unstructured or partially structured

peptides or to incorporate a metal-binding site into existing protein scaffold[97; 98; 99; 100]. A novel approach in switching system design has been taken by Woolfson *et al.*[101] Here, the authors took a step further and proposed a reversible single-sequence conformational switch triggered by zinc binding. The two target conformations were zinc finger fold and coiled coil. The choice of target conformations was dictated by the fact that the rules governing the folds are extensively studied, so the novel bio-orthogonal systems can be designed confidently. Also, both folds are independent, stable and well-defined motifs. This system formed a stable zinc finger fold but the coiled coil conformation was not as energetically favoured. An innate problem of this approach comes from the substantial constraints imposed by the duality of the sequence. The second generation of this type of switching peptide proposed by the authors aimed at improving stability of the metal-free fold – the coiled coil conformation. Thus came to life the two-component dual switching system, whereby a second peptide was designed to interact with the first – the switching one, to form a stronger coiled coil. ***Chapter 4*** describes gas-phase characterisation of this very system, and the findings are also in preparation for publication.

1.1.3 Protein-ion interactions: effect of anion from Hofmeister series

Generally, protein-salt interactions are viewed in the context of the Hofmeister series – the ranking of relative influence of ions on physical behaviour of proteins. The study of such interactions was pioneered by Franz Hofmeister[102] who observed albumin precipitation by different salts and ranked cations and anions separately on their ability to do so, with anions'

effect being more pronounced than cations'. The generally accepted order of anion series is shown in ***Figure 1-5***; ions are displayed in order of decreasing kosmotropicity. Kosmotropic ions (at the left of the figure) tend to precipitate proteins by stabilising their fold, whereas chaotropic ions (at the right of the figure) tend to solubilise proteins, often promoting denaturation. Explanation of the observed phenomenon offered by Hofmeister was based on the notion that salts either 'make' or 'break' the water structure (*i.e.* hydrogen-bonding network) around macromolecules. This notion was adhered to by the scientific community for a long time, however recent studies reveal that this effect is much more complex. Firstly, the effects of salts on proteins are more related to direct protein-ion interaction, as ions do not have long-range effect on bulk water[103]. Another factor is net charge the protein itself: if pH of the solution is above the pI of the protein, the direct Hofmeister series is observed, and if pH of solution is below the pI of the protein, the inverse series is followed (the scale at the top of ***Figure 1-5*** is reversed)[104]. And finally, salt concentration in solution plays a defining role: at higher salt concentration a direct Hofmeister series is followed, and the reverse series is characteristic of salt concentration below that[105; 106].

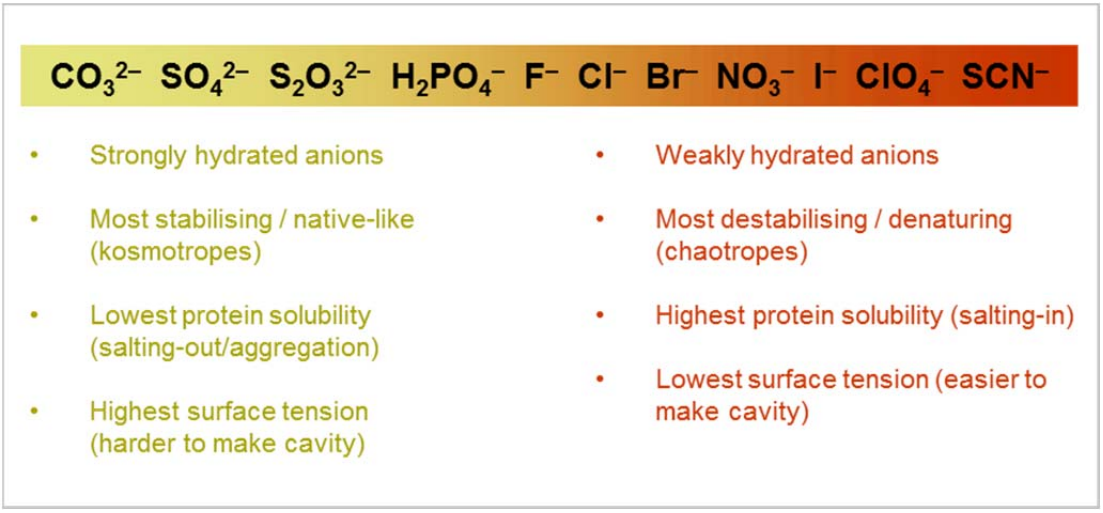


Figure 1-5 *Anion Hofmeister series. Ions on the left of the scale tend to precipitate proteins (kosmotropes), and ions on the right increase solubility of proteins (chaotropes). The chloride ion is on the border line of these two types of behaviour.*

As was demonstrated in section ***1.1.2***, interaction of metal cations with proteins by coordination bonding is a well-explored area in biology. Although it had been long established that anion binding to proteins plays a key part in many physiological and metabolic processes in the living cell, this phenomenon has been lacking scientists' attention and its role has been viewed merely as 'counter-ions' and 'co-ions'[107]. This is especially the case for the small, hard halogen ion series: F^-, Cl^-, Br^- and I^-. The fundamental impact of these anions on protein stability, allosteric propensity and catalytic activity is well-known[108; 109]. But only recently interest in a molecular mechanism of specific protein-anion interaction has re-emerged through a number of theoretical and experimental works[110; 111; 112; 113].

Zhou *et al.*[107] exhaustively surveyed crystal structures deposited in the PDB and found that more than 10,000 of halide ions are non-covalently attached to a protein surface or interior. Although an overall picture of protein-halide interaction is still lacking detail and is very fragmented, some findings were

reported by a number of research groups. For example, it was established that small anions such as F^- tend to pair with charged groups on the proteins, whereas larger anions such as I^- are more likely to interact with hydrophobic patches on the protein surface[113; 114]. Along with protein-ion interfaces, air-water interfaces have been used to model ion partitioning behaviour[109; 115]. It was found that interfacial regions are preferred by larger, less hydrated ions. The team led by Cremer[104] has confirmed earlier established finding that the protein charge and salt concentration are indeed a decisive factor whether direct or inverse Hofmeister series is observed. They also proposed a mechanism of specific chaotropic anion effect that is responsible for the hydrophobic collapse of positively charged proteins[116]. One mechanism suggests that the anion interacts with water molecules involved in hydrogen bonds with the amide of the peptide bond. The other mechanism presumes direct anion binding to the nitrogen of the amide bond causing disruption of the native-like fold and 'salting-in'.

The proteins studied in this thesis are lysozyme, cytochrome *c* and bovine pancreatic trypsin inhibitor (BPTI) (***Figure 1-6***). They all are globular proteins that have net positive charge and are well characterised by solution-, solid- and gas-phase methods: NMR and X-ray crystallography and mass spectrometry respectively. These proteins have played an important role as model systems for understanding Hofmeister effect by various methods.

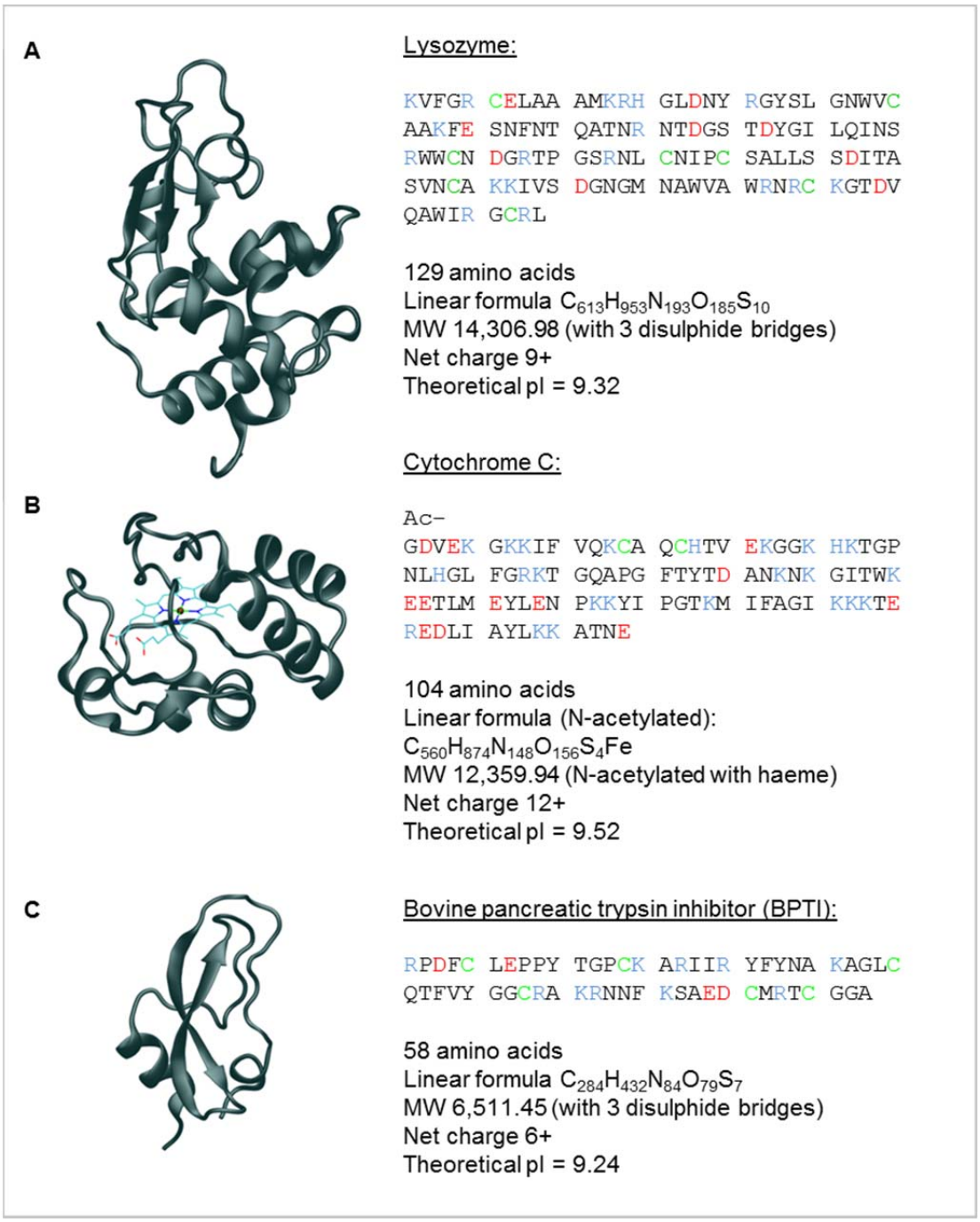


Figure 1-6 *Ribbon diagrams of proteins studied in this thesis: (**A**) lysozyme from chicken egg white (PDB 3AW6); (**B**) cytochrome c from equine heart (PDB 1HRC), featuring haeme containing Fe (protoporphyrin IX); (**C**) bovine pancreatic trypsin inhibitor (PDB 1BPI). Acidic and basic amino acid residues in protein sequences are shown in red and blue respectively. Cysteines are shown in green. Quoted pI values are calculated in: web.expasy.org/compute_pi.*

1.1.3.1 Lysozyme

The amino acid sequence of hen egg white lysozyme was first published in 1963[117] and the first X-ray structure was obtained two years later[118]. It is an enzyme whose biological role is hydrolysis of polysaccharides in bacterial cell walls, and it is found in bodily secretions performing a mild anti-bacterial function. Lysozyme contains up to four cysteine bridges (in the experiments conducted in present work, the measured protein mass suggests that on average only three of them were linked). Molecular dynamics simulations performed by Lund *et al.*[119] agreed with experimental findings in that 'salting out' in lysozyme follows the reverse Hofmeister series for pH below the isoelectric point and the direct series for pH above pI. The same authors also found that large anions are attracted to hydrophobic surfaces while smaller, well solvated ions are repelled[114]. The liquid-liquid phase transition of lysozyme was investigated by temperature gradient microfluidics under a dark field microscope by the team led by Cremer[120]. They found that in general positively charged macromolecular systems should show inverse Hofmeister behaviour only at relatively low salt concentrations, but revert to a direct Hofmeister series as the salt concentration is increased. Gokarn *et al.*[121] used effective charge measurements of hen egg white lysozyme as a direct and differential measure of ion association. They demonstrated that anions selectively and preferentially accumulate at the protein surface even at low (<100 mM) salt concentrations. The effect was dependent on anion, but not cation identity of the salt and progressively greater when binding the monovalent anions. A combination of solution- (stopped flow tryptophan fluorescence, inhibitor binding, and circular dichroism) and gas-phase studies (hydrogen exchange protection monitored by electrospray ionisation mass spectrometry) on the

effect of added salts to lysozyme refolding pathway was conducted by Kulkarni *et al.*[122] They demonstrated that the rate of formation of native lysozyme on the slow refolding pathway is significantly reduced in solutions of high ionic strength in a manner dependent on the position of the anion in the Hofmeister series. By contrast, the rate of evolution of hydrogen exchange protection monitored by ESI-MS is unchanged under the refolding conditions studied. IM-MS studies by Merenbloom *et al.*[123] of salt interactions with multiply protonated lysozyme (as well as cytochrome *c*) revealed that salts can induce compact conformations in the gas-phase ions.

1.1.3.2 Cytochrome c

Equine heart cytochrome *c* is a haeme-containing (with the haeme covalently bound[124]) protein which is a key component of the electron transport chain in mitochondria. It was first sequenced in 1962[125] and its X-ray structure was first reported in 1992[126] (earlier for bacterial cytochrome *c*). Its sequence is highly conserved across species and is a useful tool in evolutionary biology. Cytochrome *c* does not contain disulphide bridges, as the two cysteine residues and the histidine residue immediately following the second cysteine participate in haeme binding. Baglioni *et al.*[127; 128] investigated the effect of different anions of sodium salts on concentrated solutions of cytochrome *c* by small angle neutron scattering and viscosity measurements. They established that the addition of different monovalent co-ions causes the emergence of attractive interactions that follows a Hofmeister series. This phenomenon is considered as the hallmark of the gelation process (brought about by protein aggregation) promoted by specific co-ion interactions. Battistuzzi *et al.*[129] used an electrochemical and ^{1}H-NMR methods to conduct a comparative study of anion binding to bacterial cytochrome *c vs* mitochondrial (which have less

than 40% sequence similarity). They found that binding stoichiometry and strength of several anions was comparable for both and interpreted the results as indicative of the existence of common binding sites. Interestingly, these binding sites are proposed to be located in the conserved lysine-rich domain around the solvent-exposed haeme edge. This is a fine example of sequence-to-structure relationships study by 'mapping' the active site by anions. A very elegant and comprehensive study on the effect of anions on stability of gas-phase proteins and their complexes was conducted by the team led by Ruotolo[130] using IM-MS. The authors studied 6 proteins (among them cytochrome *c*) in different oligomeric states and their interaction with 12 different anions. The authors suggested that anionic components of the added salts bound to the complex either in solution or during the electrospray process. The excess energy is 'taken away' from the protein complex ion upon activation by the resulting 'shell' of counter-ions. This can result in significant overall structural stabilisation of the gas-phase protein assembly. This was the first time that the data quantified the influence of a large range of counter-ions on protein structure in the gas phase and enabled ranking of counter-ions as structure stabilizers in the absence of bulk solvent. The rank order reported in this work is considerably different when compared to the Hofmeister salt series in solution. This was an expected outcome due to the decreased influence of water on anion and protein solvation. These findings emphasise the role played by both hydration layer and protein-anion binding effects in stabilisation of macromolecules in solution. The additional benefit of the IM-MS method applied to protein-salt interaction is that the CCS obtained for the gas-phase proteins stabilised by salts can be directly compared to the coordinates obtained by NMR

spectroscopy and X-ray crystallography, as they also have salts mediating the structure.

1.1.3.3 Bovine pancreatic trypsin inhibitor

Bovine pancreatic trypsin inhibitor was sequenced in 1965[131] and its structure of was solved by 1975[132]. BPTI's function is to inhibit trypsin and other serine proteases by binding to their active site. BPTI's structure is stabilised by 3 disulphide bridges that lock its conformation, contributing to its noticeable thermal, acid and base resistivity. Partly due to its well-documented three-dimensional structure[133], BPTI has traditionally served as a model for studying protein folding. One such study aiming to investigate sequence-to-fold relationships, was testing a hypothesis that only a fraction of information contained in an amino acid sequence can be sufficient to specify a native protein structure[134]. The authors used analytical ultracentrifugation and CD and NMR spectroscopy to test this hypothesis on BPTI mutants containing many alanine residues. They found that the mutants containing up to 48% of alanine residues folded into native-like structures. Moreover, one particular mutant containing 38% of alanine residues was shown to fold into a structure very similar to that of a native protein and to retain its trypsin inhibiting function. Although BPTI is one of the models of choice for both computational[135] and gas-phase[136] studies of protein folding behaviour, the protein's interactions with Hofmeister salts is relatively unexplored. One of the few studies was concerning refolding of reduced BPTI under different conditions, and various Hofmeister salts among them[137]. The author observed direct Hofmeister series for the salts used in the study. Veesler *et al.*[138] studied BPTI in NaCl solutions with the aim of crystal growth, and they established the salt concentration threshold to be at ~1M: at concentrations

above that pre-nucleation occurred. Specific ion effects at protein surface were studied by MD simulations for BPTI (along with horseradish peroxidase) with the aim of elucidating ion adsorption at the protein surface[112]. The authors found that sulphate was always strongly attached to the proteins, choline – $(CH_3)_3N^+(CH_2)_2OH$ – showed a significant propensity for the protein surfaces, and sodium ions had a weak surface affinity, while chloride had virtually no preference for the protein surface. The simulations supported a picture of ions interacting with individual ionic and polar amino acid groups rather than with an averaged protein surface. The study highlighted the subtle differences between various salts in protein-salt interactions, emphasising the importance to employ adequate interaction models for description of such phenomena.

1.1.4 Studying protein-ion interactions in the gas phase

Traditionally, spectroscopy has been a method of choice for characterisation of biological systems and specifically protein-ion interactions. In its early days, mass spectrometry was usually only employed to obtain the weight (and perhaps purity) of the component(s) of interest[139]. However developments in soft ionisation techniques[140], and in particular in electrospray ionisation[141] have enabled researchers to analyse intact protein complexes in the gas phase, revealing stoichiometries and detailed structural traits of systems of increasing complexity[142]. When it came to analysing proteins in the presence of non-volatile salts – both metal and non-metal, mass spectrometry was very often perceived as 'incompatible' with them. As a result, such adducts were removed from protein solutions by various methods to improve desolvation efficiency and obtain 'cleaner' spectra[143; 144;

[145]. However, specific ions play a crucial role in biological systems, and their retention within the system during analysis is of paramount importance for obtaining accurate measurement of a native-like state. Over the past fifteen years, MS methods have been sufficiently refined to be successfully applied to studying weak interactions: protein-protein[142; 146], metal-protein[147; 148; 149] and salt-protein complexes[150; 151]. A wide range of parameters can be measured for such interactions: sub-unit and metal-protein stoichiometries, affinity and cooperativity of metal attachment, and oxidation state of metal ions[152].

An additional benefit of using an MS-based analysis is that it is highly complementary to traditional methods, offering analytical insight into events not accessible with spectroscopic probes. For example, in biological systems, zinc is only found as the very stable Zn^{2+} ion with d^{10} electron configuration; its saturated 3d shell is devoid of d-d transitions (***Appendix***, ***Table A 2***). Experimental conditions often do not offer the extreme conditions required for Zn^{2+} to undergo redox reactions, as charge transfer does not occur upon the ion's coordination within a protein, rendering the free and bound state of Zn^{2+} indistinguishable for UV-visible spectroscopy. Additionally, the major naturally occurring isotopes of zinc have zero nuclear spin, making it undetectable by NMR spectroscopy[153]. Thus, traditional biophysical characterisation methods are not applicable for characterisation of biological systems with zinc, and very often researchers resort to using cobalt instead for characterising zinc-containing systems[96; 154; 155].

1.2 Biological Mass Spectrometry

1.2.1 Principles and brief history of the method

Mass spectrometry (MS) is an analytical method for measuring the molecular weight of an ionised chemical compound and it has a vast range of applications in areas such as the pharmaceutical industry, medical and sports testing, defence and security, fuel and environmental analysis, manufacturing process control, space exploration and fundamental research. The basic components of a mass spectrometer include:

- an ionisation source to charge and transfer molecules into the gas phase;
- ion optics with potential gradient applied to guide and focus ions;
- a mass analyser where ions are separated based on their mass-to-charge ratio (m/z);
- a detector to register the resulting ion current.

Differential pumping is used to maintain high vacuum in the mass spectrometer necessary for mass analysis and detection (pressure is in the range of ~10^{-7} – 10^{-11} mbar at the detector, depending on instrument type). This technique owes its existence to J.J. Thomson's fundamental ion physics work at the turn of the 20th century[156]. Until 1934 mass spectrometry was only applied to inorganic analytes[157; 158]. Although the first mass spectrometer was

constructed and the first mass spectrum was obtained 99 years ago[156], the study of biological macromolecules was not feasible due to their non-volatility and thermal lability. A relatively 'soft' ionisation technique FAB (fast atom bombardment) that was first reported in 1981 by Barber *et al*.[159], enabled the study of biological molecules and was used mainly as a sequencing tool. The first FAB mass spectrum of an intact biological molecule was obtained for the doubly-charged molecular ion of insulin in 1984[160]. However, a wider application for studies of intact biological molecules received even softer ionisation techniques of ESI (electrospray ionisation) in 1968[140; 161], ESI-MS (ESI coupled to MS) in the 1980s[141], and MALDI (matrix-assisted laser desorption/ionisation) in 1985[157; 162; 163]. Both techniques brought their developers Nobel Prize in Chemistry in 2002.

1.2.2 Ionisation

For a compound of interest to be detected by mass spectrometry, it needs to be ionised. Ionisation of the analyte is the first, and probably most critical, step in mass spectrometric measurement. The breadth and relevance of analytical information obtained by mass spectrometry greatly depends on the quality of sample ionisation. This is especially crucial for biological systems, as they are extremely labile, and harsh ionisation conditions may distort them considerably. Various ionisation methods have been used in mass spectrometry in the course of the past century, however the focus here will be given to the ones used with biological systems. As mentioned just above, the two ionisation methods used for the analysis of biological macromolecules are MALDI and ESI. The former technique will be described

very briefly, as it was not used for present studies, and the latter will be given a more detailed overview.

1.2.2.1 MALDI

The MALDI technique has its origins in the laser desorption method (LD) used in surface analysis. The breakthrough that allowed large labile biological molecules to be ionised, without completely obliterating them with the laser, came when Tanaka *et al.*[163] proposed the idea of 'soft' LD (SLD). The trick here was to adjust the laser wavelength to each individual sample to avoid the energy absorption by the aromatic residues of the protein molecule, which limited the method application greatly. This method was refined by Karas *et al.*[162; 164] who first employed the use of a matrix: a small molecule mixed in with the analyte to absorb the excess energy from the laser and supply protons for ionisation. The matrix (usually cinnamic acid derivatives for peptide and protein work) is used in large excess to protein, then spotted onto the sample plate and dried. The choice of matrix is based on the fact that it should strongly absorb at the laser wavelength and be acidic to donate protons for ionisation. When the sample-matrix solid mixture is irradiated with a laser under vacuum, the matrix absorbs energy, thus shielding the analyte from the impact and sublimates together with analyte into the gas phase. The resulting ions are mainly singly-charged, with a small population of doubly-charged species and some fragments.

1.2.2.2 ESI

It took almost 20 years from the first report on electrospray ionisation (ESI) to realisation of the technique for biological molecules, however from that point

on, the 'flying elephants' (thank you Professor Fenn!)[165] of ionised biomolecules in the gas phase have been driving forward progress in biomedical science and biotechnology. ESI, along with atmospheric pressure chemical ionisation (APCI), is an atmospheric pressure ionisation (API) technique, and allows more efficient ionisation compared to low-pressure methods[157]. The principle of the electrospray process is shown in ***Figure 1-7***. The dissolved analyte under investigation is passed out of the metal capillary to which a high voltage is applied, under atmospheric pressure. Either positive or negative voltage in the order of 2 to 5 kV is applied to generate either positive or negative ions. The desolvation process is aided by heating the ESI source and the use of 'curtain' nitrogen gas. As solvent evaporates from the charged droplets containing the molecule of interest, the surface tension cannot oppose the charge repulsion forces, and the droplet forms a 'Taylor cone' which emits smaller droplets from its tip. The process repeats itself until all solvent has evaporated and the analyte ion is transferred into the gas phase[140; 166]. As mass analysis requires high vacuum, a series of focusing lenses are placed after the API source and between the series of multipole transfer optics compartments, allowing high-capacity pumps to maintain a pressure gradient along the mass spectrometer. A big advantage of the ESI method lies in the fact that it produces multiply charged ions, which allows observation of the high-molecular-weight species with mass spectrometers of much lower mass detection limit.

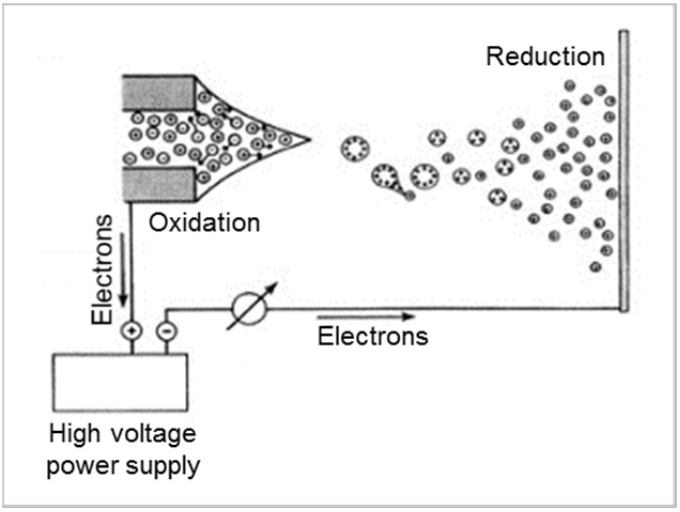


Figure 1-7 *Schematic representation of electrospray process featuring the 'Taylor cone' formation in the positive ion mode. Reprinted (adapted) with permission from Ikonomou et al.[167], copyright 1991 American Chemical Society.*

Lack of bulk protein sample has led to miniaturisation of the ESI interface – nano-ESI, or nESI[168]. Only 15 – 30 microliters of sample solution are needed for a day's experiment with flow rates of ~100 nL per hour, and sometimes a portion of that is unused and can be recovered. Due to smaller droplet size in nESI, ion generation becomes more efficient, and curtain gas is not needed for the desolvation process. Nano-electrospray also provides greater efficiency in ion introduction to the mass spectrometer due to closer proximity of the nESI tip to the inlet aperture[169]. A range of other attractive aspects of nESI include the fact that the risk of sample carryover is eliminated, it offers an alternative mechanism of ion formation due to reduced flow rates, it is tolerant to a wider variety of spray solvents and it preserves the non-covalent protein interactions[170; 171]. This has dramatically enhanced the quantity and quality of information that can be obtained by nESI-MS compared to conventional ESI and vastly expanded the area of application for this technique. ***Table 1-2*** compares principal operating parameters of conventional ESI and nano-ESI.

Parameter	ESI	Nano-ESI
Voltage	2 – 5 kV	0.6 – 2 kV
Flow rate	100 – 200 μL/min	5 – 20 nL/min
Capillary ID	~0.1 mm	1 – 10 μm
Desolvation gas	Yes	No

Table 1-2 *Comparison of principal operating parameters for ESI and nano-ESI.*

1.2.3 Mass analysis

After ions are produced, they need to be sorted according to their mass-to-charge ratio, a task performed by ion analysers. Mass analysers, just like ion sources, have evolved in a great variety of form in order to address different analytical problems. Although analyser designs, with all their ingenuity, certainly deserve to be put in the spotlight, this is however out of scope of this thesis. Keen readers may be referred to other eloquent sources, including the ones listed in the references below[157; 172; 173; 174]. The work presented in this thesis was carried out on a Q-ToF-2 mass spectrometer that employs a hybrid ion analyser: quadrupole (Q) followed by time-of-flight (ToF), and therefore some attention will be given to them here.

1.2.3.1 Quadrupole mass analyser

The concept of the quadrupole was first introduced in 1953 by Paul and Steinwedel[175], and the first commercial implementation was achieved in the mid- to late 1960s[176]. The principle of quadrupole operation is very elegant in its simplicity. A quadrupole consists of four (nearly perfectly) parallel rods, where voltages of opposite polarity are applied to each of the adjacent rods. High-frequency polarity switching on these rods is superimposed on the

constant electric field, which drive ions through the space between the rods. An ion entering the space between the rods is drawn towards the oppositely charged rod. If the polarity switching occurs before the ion discharges on that rod, the ion changes direction, and the process repeats itself until the surviving ions exit the quadrupole from the opposite end. Below is the equation[157] describing this process:

$$\Phi_0 = +(U - V \cdot \cos \omega t) \qquad \textbf{Equation 1-1}$$
$$-\Phi_0 = -(U - V \cdot \cos \omega t)$$

where Φ_0 is the potential applied to the rods (V); $\omega = 2\pi\nu$ is the angular frequency (rad/s), with ν is the RF applied to the rods; U is the direct potential (V); V is the 'zero-to-peak' amplitude of the RF voltage (V).

The process of ion filtering in the quadrupole is illustrated in ***Figure 1-8***. As ions of certain m/z value enter the quadrupole region, they experience the potentials (U and V) and frequency ω applied to the rods. As a result of this, they generate a trajectory which either compatible (resonant) with their transmission through the quadrupole or not (non-resonant). By fast scanning the RF amplitude, a range of m/z values can have a stable trajectory to be detected resulting in generation of a mass spectrum: m/z values calculated from the potential and frequency applied, and known geometry of the quadrupole. Given this ability of the quadrupole to 'filter out' certain m/z values, it is often used as a 'pre-filter' in tandem operation with other types of mass analysers. One of such types is a time-of-flight (ToF) analyser. An example of tandem experiment will be discussed later in section ***1.2.5***.

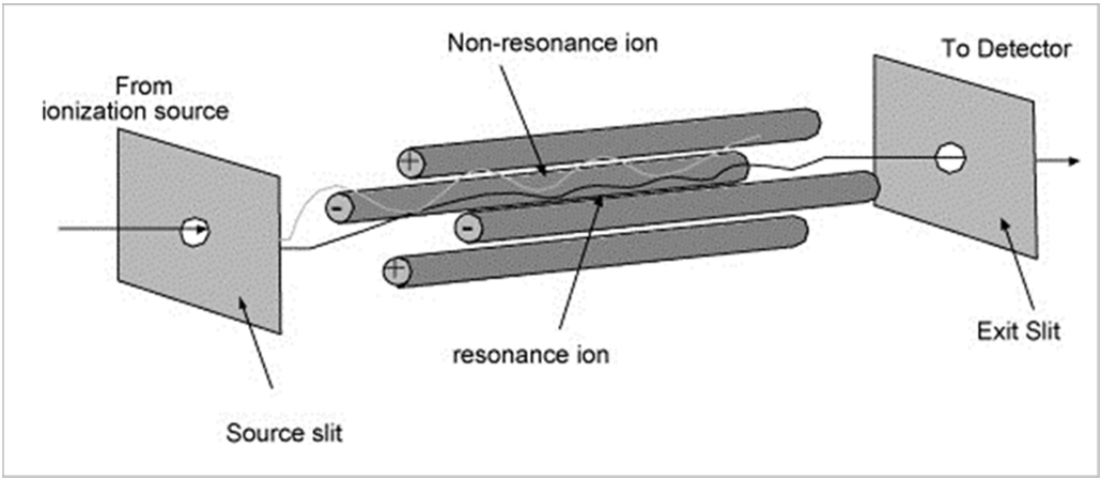


Figure 1-8 *Schematic of quadrupole mass analyser. Opposite rods have same polarity. Ions enter the quadrupole through an aperture. The ions with stable trajectory (black) are transmitted through the quadrupole and pass to the detector, whereas the ones with unstable trajectory (grey) are discharged on the rods upon collision and are not detected. Reprinted with permission from Kicman et al.[177], copyright 2007 Elsevier.*

1.2.3.2 Time-of-flight analyser

The concept of the ToF analyser was first introduced in 1946[178], and the design published in 1955 was implemented in the first commercial instrument[179]. This analyser had a linear arrangement, *i.e.* the ions were introduced at one end and detected at the opposite end (***Figure 1-9A***). ToF technology relies on pulses of ions being injected in bundles and accelerated by a potential V before entering a field-free region. The time spent by ions in this region travelling towards the detector is measured and mass-to-charge ratios are deduced from that. Thus, a charged particle of mass m attains velocity v at accelerating potential V:

$$v = \sqrt{\frac{2qV}{m}}$$ ***Equation 1-2***

where $q = ze$ is the total ion charge (z is the charge number and e is the elementary charge). Here, kinetic energy E_k is equal qV.

The time t taken by ions to travel the distance L with at velocity v between entering the field-free region and the detector is given by:

$$t = \frac{L}{v}$$ ***Equation 1-3***

By replacing v in ***Equation 1-3*** by its value from ***Equation 1-1***, it can be seen that, all else being equal, the lower the mass of an ion and the higher its charge, the faster it will travel:

$$t = L\sqrt{\frac{m}{2zeV}}$$ ***Equation 1-4***

The linear flight tube technology innately suffers from poor mass resolution, as the ions of the same m/z value are distributed along a flight path, broadening the response peak in the detector. This is schematically represented in ***Figure 1-9A*** by staggering the positions of circles representing ions of the same m/z value that have slightly different velocities. There are a number of reasons for the ions distribution along a flight path (*i.e.* time): (i) volume in which ions are formed; (ii) length of pulse during the ion injection into the flight tube; (iii) initial kinetic energy distribution during the ion injection. According to ***Equation 1-4***, mass resolution is proportional to flight time, therefore the problem can be alleviated by either decreasing the acceleration voltage or increasing the flight path. The former measure will lead to the loss of sensitivity, whereas the latter need to take into account practical consideration of the flight tube size: around 2 m flight path is needed to achieve a better resolution.

A very elegant solution to address this problem was first proposed by Mamyrin[180] in 1973: an electrostatic reflector (commonly referred to as

reflectron, and occasionally as ion mirror) was implemented at the far end of the flight tube to turn ions around and detect them next to the ion source (***Figure 1-9B***). This enabled the increase of the ion flight path two-fold, while keeping the size of the instrument within practical scale. A greater yet advantage of using the reflectron is that it corrects for the kinetic energy dispersion of ions with the same *m*/*z* values, thus increasing the instrument's resolution. As can be seen in the diagram in ***Figure 1-9B***, the slower ions (with lower kinetic energy) of the same *m*/*z* will not penetrate into the reflectron as deeply as the faster ones, and consequently will not spend as long in the reflectron as the higher-energy ones. As a result, both fast and slow ions reach the detector at the same time. This undisputable benefit, however, comes with a price of mass range limitation. Indeed, to even out the identical *m*/*z* arrival times, one is required to fine-tune the voltages, which will only cover a certain range of *m*/*z* values at a given flight tube size.

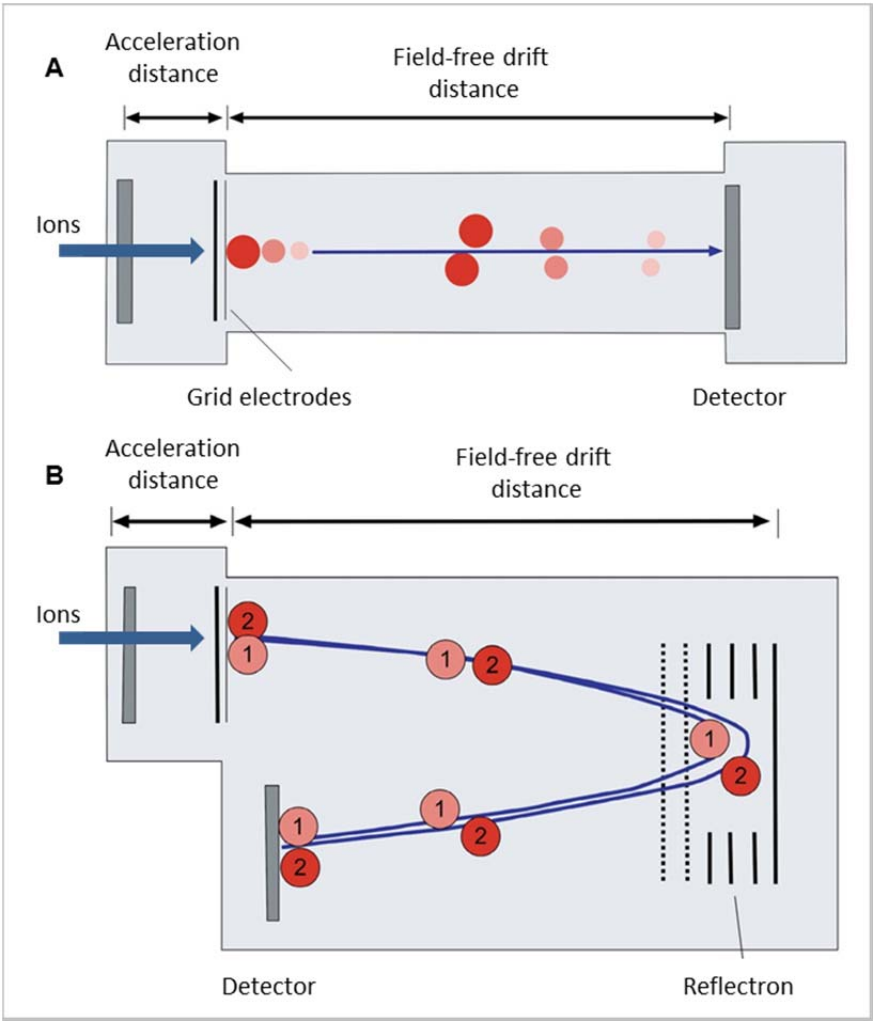


Figure 1-9 *Schematic of ToF mass analyser. (**A**) – linear geometry of the ToF mass analyser. Ions with lower m/z values are represented by smaller circles, and ions with higher m/z – larger circles. Kinetic energy dispersion is represented by staggering the circles of the same m/z. (**B**) – reflectron geometry of the ToF mass analyser. Ions of the same m/z values only are illustrated. The ions annotated '1' (pink) have lower initial kinetic energy, and the ions annotated '2' (red) have higher initial kinetic energy. Reprinted (adapted) with permission from Soares et al.*[181]*, copyright 2012 Elsevier.*

The total time spent by a charged particle in the flight tube of the reflectron geometry will be a sum of the time spent in the field-free region (moving towards the reflectron and back away from it) and the time spent in the ion mirror (from the point of entry, *via* the zero-velocity point, where the ion direction is reversed, to the point of exit). The mean velocity v of an ion entering the reflectron at initial velocity of v_i will equal $v_i/2$. The time spent

by an ion penetrating the reflectron to distance x, where it changes direction will be:

$$t = \frac{x}{v} = \frac{2x}{v_i}$$ ***Equation 1-5***

Thus the total time an ion spends in the reflectron is double this time: $t_r = 2t$. By combining this with the ***Equation 1-5***, the total time an ion spends in the reflectron can be expressed as follows:

$$t_r = 2t = \frac{4x}{v_i}$$ ***Equation 1-6***

The equations above do not take into account the fact that ions exiting the reflectron do not take exactly the same trajectory as when entering, *i.e.* the reflection takes place at a small angle. However, considering that the ion mirror is typically positioned at an angle of less than 2°, this is a good approximation of the process taking place in the reflectron.

1.2.4 Detection

The last stage of mass spectra acquisition is ion detection. Although a broad variety of detectors exist, attention will be given only to the ones that were used in the presented work: dynolite photomultiplier and microchannel plates. These two form a dual detection system of the Waters Q-ToF series instruments.

1.2.4.1 Photomultiplier detector

The first publication on photomultipliers emerged in 1935[182], and their primary applications were declared to be, among others, in facsimile transmission and movies/television (with sound!). The original authors can be readily forgiven for being oblivious to how much more their work has offered. A dynolite photomultiplier assembly can detect both positive and negative ions. It consists of two conversion dynodes (one for each particle polarity), a phosphorescent screen and a photomultiplier. Ions are accelerated towards a dynode of the opposite polarity, and secondary electrons are emitted on impact. These electrons are accelerated towards the phosphorescent screen where they are converted to photons. The photons are detected and the signal is enhanced by the dynode electrodes within the photomultiplier (with typical amplification range of 10^4 to 10^5).

1.2.4.2 Microchannel plates detector

The microchannel plate (MCP) consists of millions of very thin, conductive glass capillaries (2 to 25 μm in diameter with a centre-to-centre pitch of 3 to 32 μm[89]) micro-machined into a thin plate (***Figure 1-10A***). Each capillary or channel works as an independent continuous secondary-electron multiplier to form an array. A concept of the continuous dynode electron multiplier (as a single-channel unit) was first proposed in 1930, however the array geometry was not implemented until the 1960s[183]. Similarly to the detector type discussed above, the MCPs have their origins in image intensification devices. Parallel electrical contact to each channel is provided by the deposition of a metallic coating, usually Nichrome. As shown in ***Figure 1-10B***, when a voltage V_D is applied across the input side and output side

electrodes of the MCP, a potential gradient is built up along the channel direction. If an incident ion strikes an inner wall on the input side, it starts a cascade of electrons. These secondary electrons are accelerated by the potential gradient and travel along a parabolic path determined by the initial velocity. They then collide with the opposing wall surface, causing secondary electrons to be emitted again. In this manner, the electrons collide repeatedly within the channel as they pass towards the output side. The result is a large multiplication of the incident ion signal. Channels are often positioned at an angle (bias) to the surface of the MCP, typically within a range of 0 to 19 degrees, increasing the chances of an incident particle to strike the channel wall and be reflected at an angle to reach the opposite side of the wall. Most advanced MCP detector assemblies consist of two (or three) plates rotated at 90° relative to each other along their axes and pressed together, creating a chevron-like – v-shaped (or z-shape in case of three plates), arrangement of the channels. This tandem arrangement increases the signal gain, and the angle between the channels reduces the ion feedback. MCPs are operated at a pressure of $1 \cdot 10^{-6}$ Torr or lower to minimise background noise due to ion feedback and to prevent untimely burnout of the plates. The MCP shows high detection efficiency to electrons and ions with gains ranging from 10^4 to 10^7 depending on the MCP specification and arrangement.

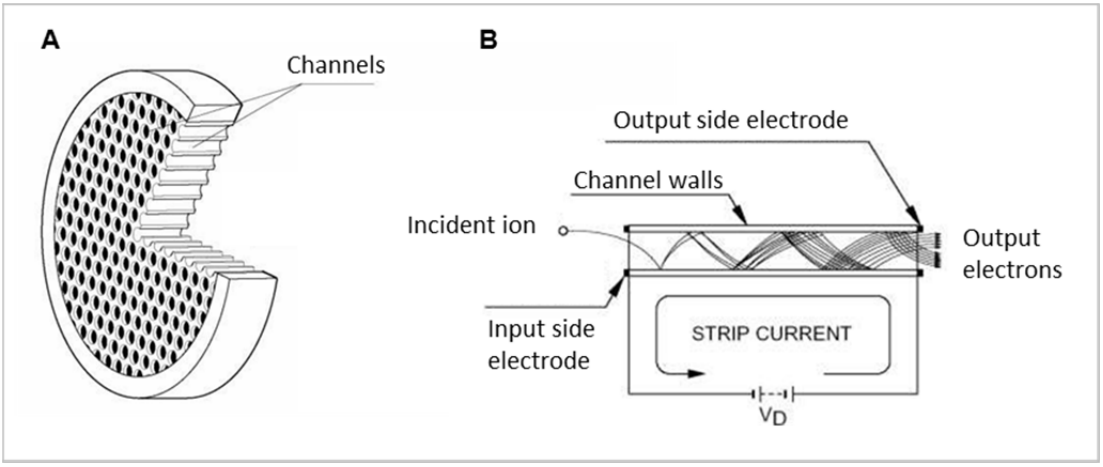


Figure 1-10 *Microchannel plate. (**A**) – MCP design featuring an array of channels in cross section. (**B**) – operating principle of an MCP channel, where incident ion starts a cascade of electrons on impact, amplifying the signal. Adapted from Hamamatsu Photonics documentation.*

1.2.5 Tandem mass spectrometry (MS/MS)

Analytical information obtained using mass spectrometry can be greatly enhanced by using tandem mass spectrometry (MS/MS), whereby intact gas-phase species (precursor ions) are selectively dissociated of into fragments (product ions) and their *m*/*z* values are measured. This process involves two or more (typically two in biological MS) stages of mass analysis with dissociation events between them. Although there is a number of ways to cause ion fragmentation, as well as technical solutions for its implementing, the focus here will be on collision-induced dissociation (CID) technique, as it was used in this work.

Selection of the precursor ion takes place in the first analyser – the quadrupole, by filtering out all other ions as explained in section ***1.2.3.1***. Ion selection can be monitored *via* the first detector (dynolite photomultiplier, discussed in section ***1.2.4.1***). CID takes place in a collision cell, where analyte

collides with inert uncharged gas molecules (typically argon at a pressure of 0.5 – 1.0 bar), and kinetic energy is partially converted into vibrational energy[184; 185], causing the protein ion to fragment, with cleavage occurring usually on the peptide bond and the sites of post-translational modifications. Fragmentation products (along with any un-fragmented ions) are passed on to the ToF analyser, where their *m*/*z* values are measured and the signal is detected on MCPs. Analysing the fragments, a valuable insight into the structure of the ion can be elucidated, *i.e.* the precursor ion is characterised according to its fragmentation pathways. This process is widely used to obtain protein structural information such as mutations, identification and localisation of post-translational modifications, non-covalent protein complexes and 3D structural information[157]. As a structure-elucidation tool, this technique was pioneered by McLafferty[186], Jennings[187] and Cooks[188], mainly in application to organic molecules, and later came into wide use for studying biopolymers, especially for protein sequencing[189] and probing the structural features of non-covalent complexes[190].

Gas-phase collisional fragmentation of proteins (or peptides) involves breaking of covalent bonds. Nomenclature of fragments due to gas-phase collisions was introduced by Roepstorff[191] and is illustrated in ***Figure 1-11***. It was proposed to class ions as either *a*, *b* or *c* if after fragmentation the charge is retained on the N-terminal fragment (presence of charge is a necessary condition for particle's detection). If the charge is retained on the C-terminal fragment, the ions are assigned as *x*, *y* or *z*. A numerical subscript is used to indicate the number of amino acid residues in the fragment. Thus, a_n and x_n ions are formed as a result of C_α-$C_{carbonyl}$ bond cleavage, b_n and y_n – $C_{carbonyl}$-N bond cleavage (the amide bond), and c_n and z_n – N-C_α bond dissociation. CID typically effects cleavage of the amide bond on the polypeptide backbone of

proteins yielding primarily *b* and *y* fragment ions (and low level of *a*-type fragment ions). Other type ions are achievable by alternative (higher-energy) fragmentation techniques.

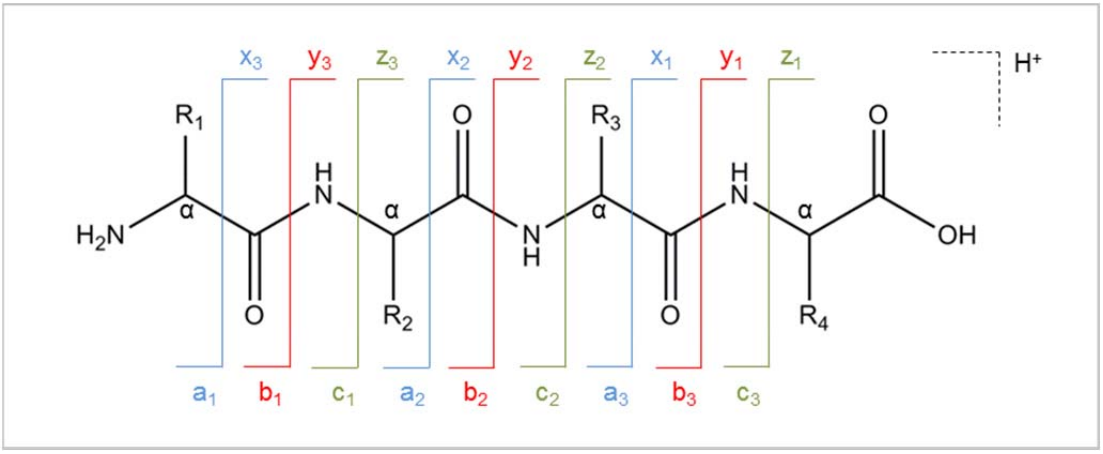


Figure 1-11 *Nomenclature of fragmentation in proteins during MS/MS experiment. Examples of dissociation techniques used to obtain relevant fragments: blue – a- and x-type, electron detachment dissociation (EDD), in negative ion mode only; red – b- and y-type, collision-induced dissociation (CID), green – c- and z-type, electron capture dissociation (ECD).*

Gas-phase collisional fragmentation of protein complexes leads to dissociation of non-covalent interactions: protein-protein or protein sub-unit associations, protein-ligand complexes, *etc*. Used alone or in combination with other techniques, both before (*e.g.* liquid chromatography, ion mobility spectrometry) and after (*e.g.* ion mobility) mass spectrometry, CID can enhance structural information that can be obtained. Apart from protein backbone fragmentation and complex dissociation, a number of additional gas-phase manipulations that can be implemented in the collision cell by activation of a protein or its complex. These include collision-induced cleaning – to remove unwanted residual solvent and salts; restructuring – collapse or spatial re-arrangement; and unfolding – sub-unit unravelling that may or may not lead to full dissociation[192].

1.3 Ion Mobility Mass Spectrometry

1.3.1 Principles and brief history of ion mobility spectrometry

Ion-mobility spectrometry (IMS) is a method for characterising chemical entities based on their gas-phase ions' transport properties in the presence of a background gas under the influence of an external electric field. A spatially 'thin' bundle of ions is pulsed into a drift cell filled with an inert buffer gas. Once in the drift tube, the ions experience a force due to the existence of an external electric field (in the case of linear ion mobility spectrometry the field decreases linearly in space, giving rise to a constant external force) inducing the ion swarm to traverse along its direction (***Figure 1-12A***). The ions undergo multiple collisions with the buffer gas molecules decelerating their progress. When the equilibrium of these two processes is reached, the ion swarm attains a constant average velocity along the electric field gradient. This velocity of the swarm is directly proportional to the electric field strength *via* a proportionality constant called the 'mobility coefficient', or simply 'mobility' (this is valid for the swarm, not individual ions). As a result, ions are separated based on their mobility, which is a function of the ions' charge, size and shape: more compact ions will traverse the cell faster than more extended ones, and so will the ions carrying higher charge than the ones with lower charge (***Figure 1-12A***). Arrival times for each ion population are registered and a plot of arrival time distribution (ATD) is reported (***Figure 1-12B***). As will be shown in section ***1.3.2***, rotationally averaged cross section for the ion-gas collision can be elucidated. IMS can be

used alone or coupled with other techniques; combined with mass spectrometry, multi-dimensional separations can be performed for structural characterisation of complex biomolecules by elucidating their size and shape. As the LDT mobility separations rely on a bundle of ions being pulsed into the separation device, this technique is very well suited for its coupling with a ToF[193; 194] technology of mass analysers, as it is pulse-triggered, and its duty cycle can be synchronised with mobility separation events.

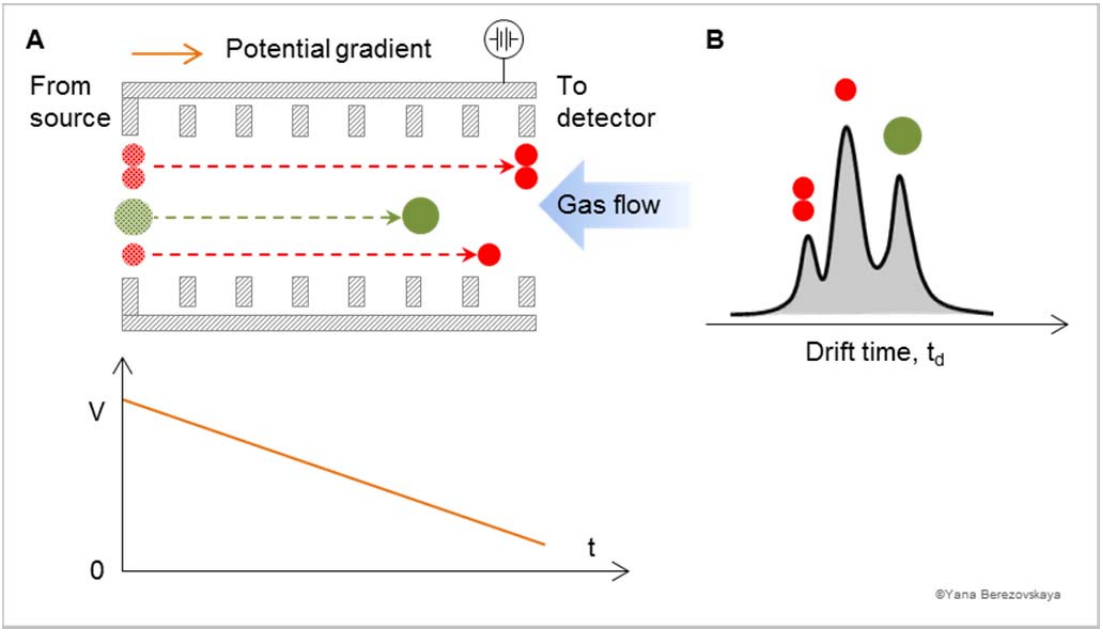


Figure 1-12 *Illustration of the principle of mobility separations in its classic implementation in a linear drift tube (LDT). (**A**) – Ion motion in the LDT during application of a linear uniform electric DC potential. The ions are transported horizontally by the electric field against the gas flow (typically helium at 2 – 4 Torr) and are shown by different colour and size circles (to symbolise different ion charge, size and shape). The red circle represents a small ion of a certain charge; the double red circles – a dimer of that ion with double its charge; the green circle represents a different larger ion. (**B**) – Arrival time distribution (ATD) plot as an output of mobility separations: higher-charged dimer arrives before its monomer; both a followed by a more extended ion.*

The theoretical foundations of the method were laid by Langevin in his works on ion-molecule associations at the beginning of the 20th century[195; 196], who recognised the collisional nature of mobility. The first prototypes of ion drift tubes were built by Cravath[197] and van de Graaff[198], and largely the same

geometry is still used today. By the 1960s IMS was widely used as a detection technique for warfare agents with only marginal excursion into the area of drugs detection[199]. Application of IMS to study biological systems coincided with the advent of ESI in mass spectrometry and, quite naturally, IMS and MS were hyphenated to enable studies of the shape and size of biomolecules and hence their tri-dimensional structure[200]. The relatively short experimental duty cycle of IM-MS allows protein dynamics to be followed on a millisecond time-scale[201].

Fundamentals of ions movement through gases are described in great detail in a book by Mason and McDaniel[202]. An array of captivating reviews exists on the subject of IM-MS of biomolecules, inclusive of instrumental development[200; 203; 204; 205; 206; 207].

1.3.2 Theory of ion mobility

Ion mobility separates species based on their migration velocity in the presence of inert gas, usually helium, under the influence of a linear electric field[200; 208]; the arrival time distribution (ATD) of ions to the detector is measured during the ion mobility experiment. The mobility (K) of an ion is determined by the velocity v_d attained under the influence of a weak linear electric field E to traverse a drift cell of length L over time t_d:

$$\frac{L}{t_d} = v_d = KE$$ **Equation 1-7**

An electric field is considered weak if the average ion energy acquired from it is small compared to the thermal energy of the buffer gas molecules. The

latter is directly affected by the buffer gas number density N (number of molecules in a volume). Therefore, the ratio E/N determines the field strength. This ratio is measured in units Td (Townsend), where 1 Td = 10^{-17} V·cm^2. A field is weak at approximate values of $E/N \leq 4$ Td; if these values are higher, mobility becomes field-dependent.

Since mobility K is dependent on N, to decouple it from experimental variables of pressure and temperature and enable data comparison, it is normalised to standard pressure (760 Torr) and temperature (273.15 K), yielding a reduced mobility K_0:

$$K_0 = K\frac{T_0 P}{T P_0}$$ ***Equation 1-8***

Mobility is determined from experimentally measured ion drift times using ***Equation 1-7***, and the rotationally averaged collision cross section Ω is calculated using the following equation:

$$\Omega = \frac{3ze}{16N}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{1}{K}$$ ***Equation 1-9***

where z is the number of charges on the ion, e [C] is the elementary charge, N [m^{-3}] is the buffer gas number density at standard conditions, μ [kg] is the reduced mass of the analyte ion and buffer gas, T [K] is the drift gas temperature, K [m^2·V^{-1}·s^{-1}] is the mobility. Derivation of this equation can be found in the book by Mason and McDaniel[202]. The reduced mass is defined by the following relation:

$$\mu \equiv \frac{m_i \cdot m_b}{m_i + m_b},$$ ***Equation 1-10***

where m_i and m_b are the masses of the analyte ion and buffer gas respectively.

As can be seen from ***Equation 1-9***, the mobility and cross section are inversely proportional to each other. Effectively, the ions are separated based on their volume and shape. Thus, elongated ions would take longer to traverse through the drift tube than their compact structural isomers.

1.3.3 Implementation of mobility separations

A number of instrumental configurations have been developed to perform separations of ions according to their mobility. The simplest instrumental set-up is the original linear drift tube (discussed above in detail and illustrated in ***Figure 1-12***), whose principles of operation rely on the theory presented in the previous section, *i.e.* the cross sections are calculated directly from drift time. This is the only 'direct', and therefore accurate, method to carry out ion mobility measurements, and it is used in the work presented in this thesis. There are two other methods that are widely used in biological work coupled to mass spectrometry: FAIMS – high-field asymmetric waveform ion mobility spectrometry, and TWIMS – travelling wave ion mobility spectrometry. A brief description of these two technologies is given below.

1.3.3.1 FAIMS

High-field asymmetric waveform ion mobility spectrometry (FAIMS)[209] is also known as 'differential mobility' spectrometry (DMS). It does *not measure* mobilities of ions being separated, but instead it *separates* them based on their mobility differences in high *vs* low field. Indeed, as was explained above, at

high field strength the mobility becomes field-dependent (typical planar FAIMS device is operated at ~70 – 80 Td[210]). A good analogy to compare FAIMS to the classic LDT was given by Guevremont *et al.*[211]: "FAIMS resembles a quadrupole analyser, whereas DT-IMS is a time-of-flight analyser". Here, ions drift either between two parallel planar electrodes or in the space between two concentric cylinders (the electrode sheets are 'wrapped' onto themselves). The space between the electrodes is filled with buffer gas flowing in the same direction as the ions are passed, pushing them through the gap (***Figure 1-13***). One of the electrodes is kept at ground potential, and an alternating electric field (asymmetric waveform) is applied to the other electrode perpendicular to the gas flow. The asymmetric waveform consists of two components: a high-voltage short-term component and a lower-voltage longer-term component of opposite polarity[212]. As a result, ions alternately experience two distinct field strengths and oscillate between the electrodes preferentially moving towards one of them. To prevent all the ions from discharging on electrodes and enable their transmission, a compensation voltage (CV) is applied to correct their trajectory. Different CVs are required to transmit ions of different size/shape and charge through the FAIMS cell. Analogous to the approach applied in the quadrupole mass filters, CV can be either scanned across a range of values to transmit ions in the bracket of interest, or fixed at a specific value to select a discrete ion population. As the difference in mobilities in high and low fields is reflected in the value of CV, this is the value that is reported for FAIMS separations (*cf.* with drift time t_d for LDT)[213].

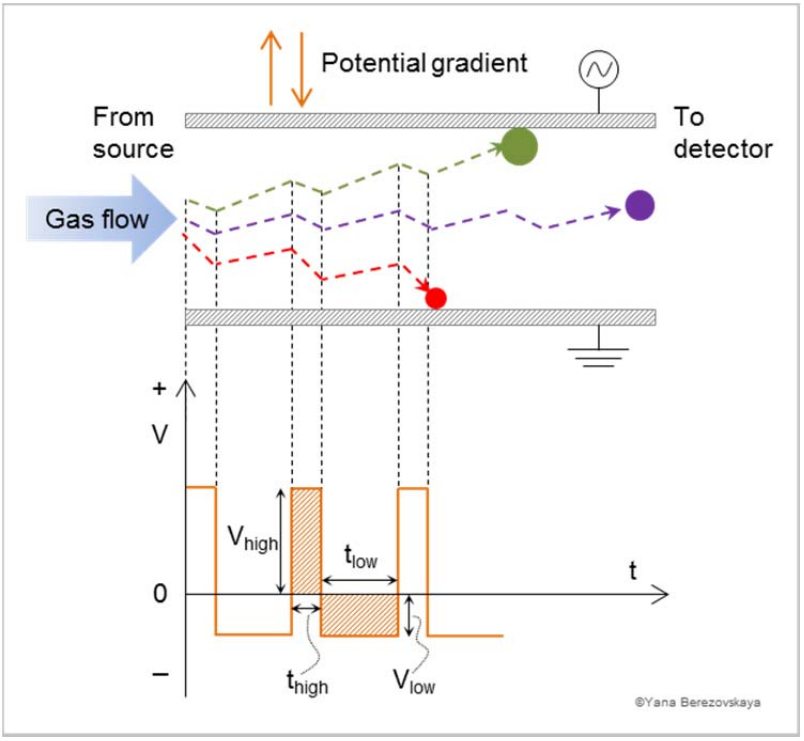


Figure 1-13 *Illustration of the ion motion between two parallel FAIMS plates during application of an electric potential V(t). The ions are transported horizontally by a gas flow and are shown by different colour and size circles (to symbolise differential mobilities as a function of a range of their properties). The voltage-time product (orange shaded areas on the positive and negative portions of the waveform) at high and low potentials are equal:* $V_{high} \cdot t_{high} = V_{low} \cdot t_{low}$.

One of the advantages of using FAIMS over LDT-IMS is its applicability at atmospheric pressures and ambient temperatures in continuous-flow regimes, enabling a FAIMS cell to be modularly attached to a mass spectrometer between the ion source and MS entrance aperture. This allows the sample to be 'pre-cleaned' before mass analysis, in analogy with liquid chromatography, and has been implemented commercially[214; 215]. Another advantage of FAIMS is better sensitivity, as it is a continuous-flow technique with improved duty cycle, and the ions are better focussed between the electrodes. Among the disadvantages is the use of high fields that can cause unfolding of labile biological ions, and there is no direct correlation between the apparent mobility of an ion and its cross section[216].

1.3.3.2 TWIMS

Travelling wave ion mobility spectrometry (TWIMS) is the commercial answer to the problem of low IM resolution of classic ion mobility separations[217] developed at Waters Corporation. The TWIMS drift device is fully integrated into a vacuum chamber of a ToF mass spectrometer and is marketed as Synapt®. In this sense the travelling wave device is comparable to the LDT, operating at pressures lower than 1 atm of buffer gas (typically nitrogen at 0.2 – 0.5 mbar), it requires an ion pulse to start a mobility separation, but is not capable of producing ion collision cross sections as output (nevertheless, calibration of drift-times to those measured by LTD for particular ions have been used to yield approximate cross sections in helium from these measurements performed in nitrogen). The drift chamber of a TWIMS device is a stacked ring ion guide (SRIG) capable of transmitting a travelling wave potential. The SRIG used for TWIMS is the central one from an array of three present in a Synapt®: the first one being used to accumulate ions prior to injection into the mobility device, and the last one for transmitting the separated ions into the mass analyser (ToF). The operation of the SRIG is illustrated in ***Figure 1-14***, whereby opposite phases of the RF voltage are applied to consecutive ring electrodes, producing a radially-confining potential. This potential prevents ions from radial diffusion from their intended path, rendering a highly focused beam. A transient DC potential pulse is superimposed on the RF and is applied to sequential electrodes producing an electric field propelling the ions along the axis of the SRIG. The series of potential hills and valleys are subsequently applied to the next pair of electrodes at regular intervals creating a travelling wave, inducing, on average, ion motion in the direction of wave propagation. As ion motion is impeded by the frictional forces of the buffer gas, ions with a

higher mobility will be affected by this process to a lesser extent, giving rise to shorter drift times. For large wave heights (which depends on the instrument settings and ion charge), all ions will be carried to the end of the drift tube by a single wave, and no separation will occur. Only at sufficiently low wave heights will ions 'roll over' the waves.

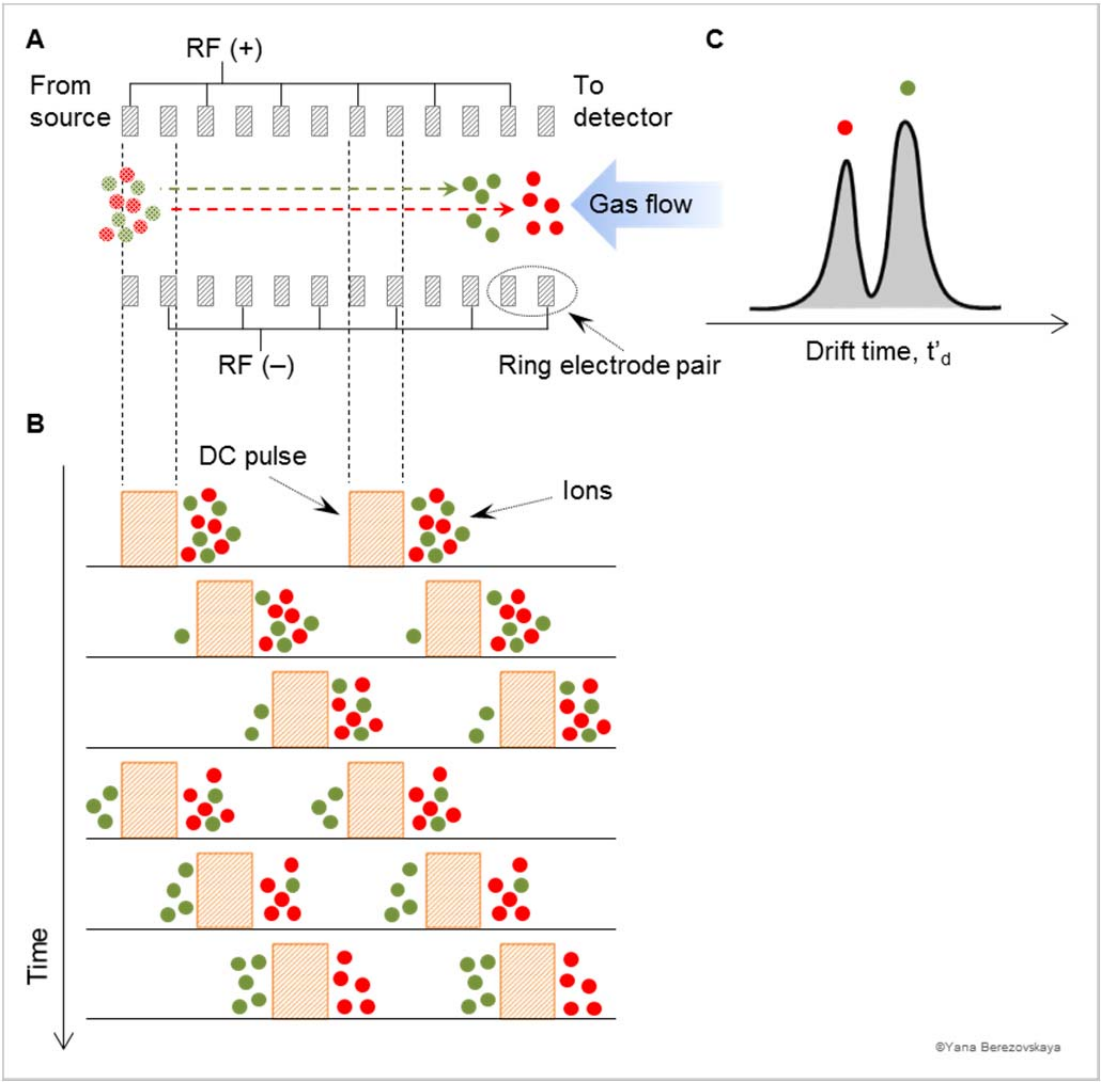


Figure 1-14 *Illustration of the principle of mobility separations in a travelling wave ion guide (TWIG). (**A**) – Schematic of the stacked ring ion guide (SRIG) device featuring ring electrode pairs. (**B**) – Ion motion in the SRIG during application of a T-wave. The ions are transported horizontally by the transient DC potential against the gas flow (red – high-mobility ions, green – low mobility ions). (**C**) – Arrival time distribution (ATD) plot as an output, illustrating apparent drift time t'_d before calibration.*

A fundamental difference between the LDT-IMS and TWIMS is that in the latter case there is no direct linear proportionality between the ion drift time and its collision cross section[218]. To overcome this, all drift time data obtained on the T-wave need to be calibrated to the data obtained on the LD device[219], for which a formal database is maintained by the Clemmer research group[99]. A full understanding of the effect of the T-wave device on the labile biological ions is currently being accumulated, however it must be pointed out that certain amount of heating-induced conformational changes have been observed in proteins[220].

1.3.4 Peptide and protein structural elucidation by ion mobility mass spectrometry

Mass spectrometry, even as a stand-alone technique, can provide useful insights into protein conformational changes in the gas phase. The first such work was carried out more than 20 years ago[221]. Here, the authors presented bovine cytochrome *c* with different pH conditions and observed the change in charge state distribution. Thus, the number of available ionisable sites on the protein was used as a measure of 'unfoldedness' of the conformation. MS-based approaches offer a range of additional tools to interrogate the system under investigation. For example, collision induced dissociation can be applied to interrogate the arrangement of sub-units in the protein complex and the strength of interaction between them[190; 222; 223]. ESI-MS is easily coupled to liquid-phase techniques, such as chromatography, which enables the study of biochemical processes in online arrangement[224]. High sensitivity (down to femtomolar concentrations), dynamic range (~10^6) and selectivity of MS also allows monitoring of many/all components at once[225]. The gentle

process of nESI preserves non-covalent interactions, and protein-ligand and protein-protein interactions can be studied. Many biochemical processes reach equilibration within a few seconds after initiation, which renders monitoring of such processes before the equilibrium is established inaccessible by X-ray crystallography and NMR spectroscopy. The millisecond time-scales of MS-based techniques offer an attractive alternative to study non-equilibrium states of protein systems and fast dynamical processes[226].

It has also become possible to determine conformations of proteins and their complexes using IM-MS[200; 227; 228]. It allows users to calculate collision cross sections and hence obtain valuable information on the conformations adopted in the gas phase[136; 200; 203]. Ion mobility adds an extra dimension to mass spectrometry, as it can resolve structural isomers[229], probe non-covalent interactions[227; 228; 230] and monitor catalytic activity[231]. Ion mobility data can be interpreted with the help of molecular modelling, whereby theoretical collision cross section of a protein can be compared to experimental values. Based on this comparison, candidate geometries can be proposed. Ion mobility mass spectrometry has proven to be a very appropriate technique to study sequence-to-structure relationships, since the solvent effects are absent, and the conformations observed *in vacuo* are defined mainly by the sequence[232; 233]. These MS-based approaches to determining structure represent an attractive complementary rapid route to analysis, assessing whether a given design strategy has been successful in obtaining the correct fold. Along with several traditional analytical biophysical characterisation methods, IM-MS might form part of a protein design platform.

The past 30 years have seen growth of molecular weight that has and can be examined by mass spectrometric measurements. For example a group led by Gary Siuzdak measured the mass of an intact tobacco mosaic virus (TMV) using a ToF instrument[234]. The capsid of TMV is a multi-protein assembly consisting of approximately 2140 identical proteins, and its calculated molecular weight is ~40.5 MDa. These new intact protein molecular weight frontiers reached by MS have now far surpassed those of X-ray crystallography and NMR spectroscopy, the latter being traditionally confined to small soluble proteins only.

1.4 Complementary biophysical techniques to study peptides and proteins

Each method available to scientists to characterise protein systems has its strengths and weaknesses, and very often only a wise combination of them can deliver reliable and informative results. Most common biophysical techniques that are used to characterise proteins and their interactions are briefly overviewed here.

1.4.1 Solution-phase techniques

1.4.1.1 NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy provides structural determination at atomic resolution. NMR is a non-destructive solution-phase

technique, which is considered a 'natural' environment for proteins. It employs magnetic properties of atomic nuclei positioned in close proximity of each other that have different nuclear spins. The atoms may be in proximity either a result of being adjacent along the polypeptide backbone or due to three-dimensional fold of the protein[235]. NMR relies on the absorption of radiofrequency energy by nuclei within a strong magnetic field. Since nuclei within a given molecule have different bonding arrangements or neighbouring atoms, each nucleus gives rise to a characteristic resonant frequency. In biological NMR, the identification of individual proteins uses molecules enriched with ^{13}C and ^{15}N to allow assignment. By analysing this information against the constraints of feasible distances, angles and torsions between the atoms, a 3D model of a protein can be reconstructed, along with their regions and surfaces of interaction with other molecules, binding affinities and conformational dynamics over a wide timescale of picoseconds to seconds. Monitoring of dynamic processes, however, only reveals the average state. Currently, the size of biological complexes investigated by solution-state NMR typically does not exceed 100 kDa for atomic resolution[236]. A major drawback of this method is the amount of sample that is required for analysis (albeit recoverable): the typically required amounts of 0.5 mL at ~1 mM concentration can be difficult to obtain for proteins. Another limiting factor is the time needed for solving the structure.

1.4.1.2 UV-visible spectroscopy

UV-visible spectroscopy is a low-resolution technique that is useful for identifying and monitoring changes in protein structure triggered by an external stimulus. The process relies on absorption of energy in the UV-visible spectral region by π-electrons and non-bonding electrons resulting in

their excitation from the ground state to higher anti-bonding molecular orbitals. Absorption of light is measured as a function of its frequency or wavelength to obtain a spectrum. Absorption and wavelength (or frequency) depends on the chemical nature and the local environment of the chromophore. In application to proteins, the process relies on their intrinsic absorbance of UV light by tryptophan residues (with small contribution from tyrosine and phenylalanine)[237]. Hence it can be used as a probe for the conformational state of the protein, as well as its coordination geometry and extent of metal binding[154]. Among the advantages of this method is its speediness, ease of implementation, relatively low cost, and low sample amounts are required (and they are recoverable): ~70 μL at ~10^{-5} M, with no limit imposed on the size of macromolecule. A significant disadvantage of UV-visible spectroscopy is that it is not a universal method: it is not applicable to compounds that do not have chromophore, and some molecules are spectroscopically silent, as was explained in section ***1.1.4***.

1.4.1.3 Circular dichroism

Circular dichroism (CD) spectroscopy measures the differences in absorption between left and right circularly polarised light (UV and visible) when the chromophore is chiral. Two types of chromophore chirality are found in biomolecules: intrinsic or due to being linked to chiral centre, and due to asymmetric three-dimensional environment[238]. CD is a very versatile method that yields complementary structural information from different spectral regions. It encompasses secondary structure and folding motifs; monitoring of conformational changes; thermal stability investigations; integrity of co-factor binding sites; metal binding (can resolve individual d–d electronic transitions as separate bands[239]). The advantages of this method include low

sample volume requirements (down to ~3 μL) at the same concentration ranges as for UV-visible spectroscopy (~10^{-5} M), no macromolecule size limit, and sample can be investigated under variety of conditions including physiological. Among the drawbacks of the method are restrictions on solvents (chlorinated solvents cannot be used) and salt concentrations (high salt concentrations should be avoided: ≤5 mM). A notable drawback is low structural resolution that cannot be used for quantitation purposes.

1.4.1.4 Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) is another versatile tool to characterise proteins based on their sedimentation velocity/equilibrium in the centrifugal field, monitored by absorbance measurements. Analytical information available with this method include examination of sample purity, determination of molecular weight, analysis of conformational changes and associating systems[240]. Method of sedimentation equilibrium specifically pertains to association studies, whereby a range of metrics are yielded: molecular weights of each monomer before association and of complex, as well as stoichiometry of heterogeneous components and strength of interaction between them. A wide range of macromolecule sizes is covered by AUC: from 100 Da to 10 GDa. Sample concentrations required for AUC are in the same range as for the two previously described methods – UV-visible and CD spectroscopies (~10^{-5} M). A significant drawback of the method is tremendous hazard associated with a risk of catastrophic failure due to rotational kinetic energy of the rotor during operation.

All spectroscopic measurements in this thesis were performed by Dr Craig Armstrong at the University of Bristol, and some data can be found in ***Appendix***.

1.4.2 Solid-phase technique: X-ray crystallography

X-ray crystallography (XRC) is a sub-atomic resolution (better than 0.85 Å, achievable on synchrotron beam-lines)[241] technique that has no rivals in the level of structural information attainable and is considered the 'gold standard'. This technique studies the proteins and their complexes in the solid state, which renders it unsuitable for any dynamic processes characterisation. Also, it relies on very high quality crystals of proteins, and obtaining the crystals is a laborious time-consuming process of trial-and-error, also requiring substantial amounts of sample. Very often it is impossible to obtain crystals at all. When these obstacles are surmounted, the crystal can be interrogated by diffraction of X-rays by electron clouds surrounding all atoms in the crystal lattice. The resulting electron density map is deconvoluted to yield atomic distances and the three-dimensional architecture in the molecule. Data analysis is another time-consuming aspect of XRC, however this is well rewarded by the fact that there is no size limit imposed on macromolecular complexes under investigation[242].

1.4.3 Computational technique: molecular modelling

Molecular dynamics (MD) simulations provide a very useful method to complement IM-MS data, as detailed structural insight into protein

behaviour can be obtained. MD-generated candidate geometries can be used to 'fill-in' experimental cross sections and used for evaluating the data. Protein behaviour can be modelled either in solution (explicit or implicit) or *in vacuo*. MD enables the study of complex dynamic processes, including conformational changes, protein folding, molecular recognition and binding, and ion transport. The output of MD simulations is a set of coordinates of the modelled structure, which can be compared with IM-MS data. This set of coordinates is then used to calculate theoretical CCS, for which three different methods are used (implemented in MobCal program[243; 244]): the projection approximation (PA), the exact hard sphere scattering (EHSS), and the trajectory method (TM). In PA method, the ion is modelled by a collection of overlapping hard spheres with radii equal to hard sphere collision distances[245]. The orientationally-averaged geometric cross section is determined by averaging the geometric cross section over all possible collision geometries. This method only works well if molecular surfaces are convex. Real-life systems, however, contain cavities which are either shielded from collisions by other parts of the macromolecule, or multiple encounters between the ion and buffer gas can take place within a cavity due to reflections. Adjustments were made to this method to account for such events, which gave rise to the EHSS model[244]. Here, the ion is modelled in the same way as for PA. However the orientationally-averaged momentum transfer cross section is calculated by determining the scattering angles between the incoming buffer gas atom trajectory and the departing buffer gas atom trajectory. TM treats the ion as a collection of atoms, each one represented by a potential[243]. The effective potential is obtained by summing over the individual atomic contributions and then trajectories are run in this potential to obtain the scattering angle (the angle between the incoming and

departing buffer gas atom trajectory). The orientationally-averaged collision integral is determined by averaging over all possible collision geometries. The trajectory method, which accounts for long-range interactions, gives the most reliable estimate. The PA method generally gives numbers that are lower than those obtained by TM, while EHSS – higher[200]. This trend is more pronounced for larger systems.

A major limitation of MD simulations is the short time-scales accessible by this method – typically nanoseconds[246], whereas IM-MS experiments deal with much longer time-scales of a few milliseconds, during which time structural rearrangements can take place. For comparison, the simplest structural elements of proteins, such as disordered loops, α-helices and β-hairpins take longer to form: 0.007 – 0.09 μs, 0.4 – 2.0 μs and 0.8 – ~20 μs respectively[12]. Folding of most proteins to their native state is usually complete within a second[247]. Continuous improvements are being made to enhance sampling methods to reduce computational time, and improved correlation between experimental and theoretical results is anticipated in the future[248; 249].

All MD calculations in this thesis were performed by Dr Massimiliano Porrini at the University of Edinburgh, and some experimental details can be found in ***Appendix***.

1.5 Contribution of this work

This work aims to provide an insight into protein-ion interactions using a gas-phase platform. Behaviours considered here include conformational changes of proteins and peptides in response to ion binding and temperature, as well as specificity and strength of such binding.

Details of experimental approach are presented in ***Chapter 2***. Described here are particulars of mass spectrometry and ion mobility mass spectrometry approach in application to biological samples. A workflow of IM-MS experiment is explained, inclusive of molecular dynamics simulations that are used to complement experimental findings.

Chapter 3 describes a gas-phase method development quest to characterise synthetic peptide-based system in an attempt to understand how a protein sequence dictates its fold. The system studied is a consensus zinc finger sequence (vCP1) that responds to zinc ion binding by adopting a zinc finger fold. This study was used to optimise characterisation parameters for analysis of a more complex peptide-based switching system presented in the following chapter. Stoichiometry of binding is probed, peptide-metal and peptide-peptide affinities are quantified by titration. Gas-phase conformations are evaluated by IM-MS and compared to those by MD simulations.

Chapter 4 describes refinement of the gas-phase analytical platform discussed in the preceding chapter in application to the peptide-based dual switching system (ZiCop) that is capable of adopting either of the two conformations in response to a stimulus: zinc finger or coiled coil. The zinc-

bound state of the switch is very similar to the one described using the example of vCP1. The coiled coil state involves the peptide's interaction with another peptide to form a stable two-helix parallel dimeric coiled coil. Stoichiometry of binding is probed, peptide-metal and peptide-peptide affinities are quantified by titration and CID. Gas-phase conformations are evaluated by IM-MS and compared to those by MD simulations.

Chapter 5 investigates protein-salt interactions using three proteins (lysozyme, cytochrome *c* and BPTI) and one salt (iodide anion). Conformational changes that take place in response to HI adductation and temperature are discussed. Experimental collision cross sections of apo-proteins, as well as iodide-bound with up to three anions, are obtained and compared to those of crystal structures in the absence of any ions. Qualitative conclusions are drawn.

Chapter 6 concludes the findings of this thesis and briefly outlines the context of potential further developments in the area of protein-ion interactions.

1.6 References

1. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. (1983). *Molecular Biology of the Cell*, Garland Publ., Inc.

2. Stryer, L. (1995). *Biochemistry*. Fourth Edition edit, W.H.Freeman and Company, New York.

3. Dobson, C. M., Sali, A. & Karplus, M. (1998). Protein folding: a perspective from theory and experiment. *Angew. Chem., Int. Ed.* **37**, 869-893.

4. Despa, F., Orgill, D. P. & Lee, R. C. (2005). Molecular crowding effects on protein stability. *Ann. N. Y. Acad. Sci.* **1066**, 54-66.

5. Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L. & Hallewell, R. A. (1992). Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature (London)* **358**, 347-51.

6. Linse, S., Brodin, P., Johansson, C., Thulin, E., Grundstrom, T. & Forsen, S. (1988). The role of protein surface charges in ion binding. *Nature* **335**, 651-652.

7. Head-Gordon, T. & Brooks, C. L. (1987). The role of electrostatics in the binding of small ligands to enzymes. *J. Phys. Chem.* **91**, 3342-3349.

8. Linse, S., Joensson, B. & Chazin, W. J. (1995). The effect of protein concentration on ion binding. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 4748-52.

9. Anfinsen, C. B., Haber, E., Sela, M. & White, F. H., Jr. (1961). Kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1309-14.

10. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* **181**, 223-30.

11. Fetrow, J. S., Zehfus, M. H. & Rose, G. D. (1988). Protein folding: new twists. *Nat. Biotechnol.* **6**, 167-71.

12. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14**, 76-88.

13. Channon, K., Bromley, E. H. C. & Woolfson, D. N. (2008). Synthetic biology through biomolecular design and engineering. *Curr. Opin. Struct. Biol.* **18**, 491-498.

14. Lu, Y. (2010). Metal ions as matchmakers for proteins. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 1811-1812.

15. Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. & Thornton, J. M. (2008). Metal ions in biological catalysis: from enzyme databases to general principles. *JBIC, J. Biol. Inorg. Chem.* **13**, 1205-1218.

16. Ismail, S. A., Vetter, I. R., Sot, B. & Wittinghofer, A. (2010). The structure of an Arf-ArfGAP complex reveals a Ca2+ regulatory mechanism. *Cell (Cambridge, MA, U. S.)* **141**, 812-821.

17. Berg, J. M. & Shi, Y. (1996). The galvanization of biology: a growing appreciation for the roles of zinc. *Science (Washington, D. C.)* **271**, 1081-5.

18. Klee, C. B. & Vanaman, T. C. (1982). Calmodulin. *Adv. Protein Chem.* **35**, 213-321.

19. Waldron, K. J. & Robinson, N. J. (2009). How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **7**, 25-35.

20. Lindskog, S. (1997). Structure and mechanism of carbonic anhydrase. *Pharmacol. Ther.* **74**, 1-20.

21. Shirran, S. L. & Barran, P. E. (2009). The Use of ESI-MS to Probe the Binding of Divalent Cations to Calmodulin. *J. Am. Soc. Mass. Spectrom.* **20**, 1159-1171.

22. Rosenzweig, A. C. (2002). Metallochaperones bind and deliver. *Chem. Biol.* **9**, 673-677.

23. Goyal, K. & Mande, S. C. (2008). Exploiting 3D structural templates for detection of metal-binding sites in protein structures. *Proteins: Struct., Funct., Bioinf.* **70**, 1206-1218.

24. Andreini, C., Bertini, I., Cavallaro, G., Najmanovich, R. J. & Thornton, J. M. (2009). Structural Analysis of Metal Sites in Proteins: Non-heme Iron Sites as a Case Study. *J. Mol. Biol.* **388**, 356-380.

25. Stoker, H. S. (2010). *General, Organic, and Biological Chemistry*. Fifth Edition edit, Brooks/Cole.

26. Chasapis, C. T., Loutsidou, A. C., Spiliopoulou, C. A. & Stefanidou, M. E. (2012). Zinc and human health: an update. *Arch. Toxicol.* **86**, 521-534.

27. Pearlmutter, A. F. & Soloff, M. S. (1979). Characterization of the metal ion requirement for oxytocin-receptor interaction in rat mammary gland membranes. *J. Biol. Chem.* **254**, 3899-906.

28. Laity, J. H., Lee, B. M. & Wright, P. E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **11**, 39-46.

29. Armstrong, C. T., Boyle, A. L., Bromley, E. H. C., Mahmoud, Z. N., Smith, L., Thomson, A. R. & Woolfson, D. N. (2009). Rational design of peptide-based building blocks for nanoscience and synthetic biology. *Faraday Discuss.* **143**, 305-317.

30. Pessi, A., Bianchi, E., Crameri, A., Venturini, S., Tramontano, A. & Sollazzo, M. (1993). A designed metal-binding protein with a novel fold. *Nature* **362**, 367-369.

31. Wilson, G. S. & Hu, Y. (2000). Enzyme-Based Biosensors for in Vivo Measurements. *Chem. Rev. (Washington, D. C.)* **100**, 2693-2704.

32. Smith, R. G., D'Souza, N. & Nicklin, S. (2008). A review of biosensors and biologically-inspired systems for explosives detection. *Analyst (Cambridge, U. K.)* **133**, 571-584.

33. Anfinsen, C. B., Redfield, R. R., Choate, W. L., Page, J. & Carroll, W. R. (1954). Studies on the gross structure, cross-linkages, and terminal sequences in ribonuclease. *J Biol Chem* **207**, 201-10.

34. Chan, H. S. & Dill, K. A. (1993). The protein-folding problem. *Phys. Today* **46**, 24-32.

35. Yue, K. & Dill, K. A. (1992). Inverse protein folding problem - designing polymer sequences. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4163-4167.

36. Nomura, A. & Sugiura, Y. (2004). Sequence-Selective and Hydrolytic Cleavage of DNA by Zinc Finger Mutants. *J. Am. Chem. Soc.* **126**, 15374-15375.

37. Papworth, M., Kolasinska, P. & Minczuk, M. (2006). Designer zinc-finger proteins and their applications. *Gene* **366**, 27-38.

38. Woolfson, D. N. (2005). The design of coiled-coil structures and assemblies. *Adv. Protein Chem.* **70**, 79-112.

39. Ambroggio, X. I. & Kuhlman, B. (2006). Design of protein conformational switches. *Curr. Opin. Struct. Biol.* **16**, 525-530.

40. Engelke, D. R., Ng, S.-Y., Shastry, B. S. & Roeder, R. G. (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* **19**, 717-728.

41. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

42. Wolfe, S. A., Nekludova, L. & Pabo, C. O. (2000). DNA Recognition by Cys2His2 Zinc Finger Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 183-212.

43. Negi, S., Imanishi, M., Matsumoto, M. & Sugiura, Y. (2008). New Redesigned Zinc-Finger Proteins: Design Strategy and Its Application. *Chem. Eur. J.* **14**, 3236-3249.

44. Iuchi, S. & Kuldell, N. (2005). *Zinc Finger Proteins: From Atomic Contact to Cellular Function*, Plenum Publishers.

45. Krishna, S. S., Majumdar, I. & Grishin, N. V. (2003). Structural classification of zinc fingers: survey and summary. *Nucl. Acids Res.* **31**, 532-550.

46. Brown, D. (1984). The role of stable complexes that repress and activate eukaryotic genes. *Cell* **37**, 359-365.

47. Miller, J., McLachlan, A. D. & Klug, A. (1985). Repetitive Zinc-Binding Domains in the Protein Transcription Factor Iiia from Xenopus Oocytes. *EMBO J.* **4**, 1609-1614.

48. Li, W., Zhang, J., Wang, J. & Wang, W. (2008). Metal-Coupled Folding of Cys2His2 Zinc-Finger. *J. Am. Chem. Soc.* **130**, 892-900.

49. Berg, J. M. (1990). Zinc Finger Domains: Hypotheses and Current Knowledge. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 405-421.

50. Parraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Elton, T. Y. & Klevit, R. E. (1988). Zinc-Dependent Structure of a Single-Finger Domain of Yeast ADR1. *Science* **241**, 1489-1492.

51. Klug, A. (2010). The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annu. Rev. Biochem.* **79**, 213-231.

52. Fox, A. H., Liew, C., Holmes, M., Kowalski, K., Mackay, J. & Crossley, M. (1999). Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *EMBO J.* **18**, 2812-2822.

53. Grishin, N. V. (2001). Treble clef finger - a functionally diverse zinc-binding structural motif. *Nucleic Acids Res.* **29**, 1703-1714.

54. Schwabe, J. W. R., Chapman, L., Finch, J. T. & Rhodes, D. (1993). The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell (Cambridge, Mass.)* **75**, 567-78.

55. Christy, B. & Nathans, D. (1989). DNA Binding Site of the Growth Factor-Inducible Protein Zif268. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8737-8741.

56. Pavletich, N. P. & Pabo, C. O. (1991). Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 Å. *Science* **252**, 809-817.

57. Grove, T. Z., Cortajarena, A. L. & Regan, L. (2008). Ligand binding by repeat proteins: natural and designed. *Curr. Opin. Struct. Biol.* **18**, 507-515.

58. Regan, L. & Clarke, N. D. (1990). A tetrahedral zinc(II)-binding site introduced into a designed protein. *Biochemistry* **29**, 10878-10883.

59. Thorp, H. H. (1998). Bioinorganic chemistry and drug design: here comes zinc again. *Chem. Biol.* **5**, R125-R127.

60. Paillard, G., Deremble, C. & Lavery, R. (2004). Looking into DNA recognition: zinc finger binding specificity. *Nucleic Acids Res.* **32**, 6673-6682.

61. Kim, C. A. & Berg, J. M. (1996). A 2.2 A resolution crystal structure of a designed zinc finger protein bound to DNA. Nat. Struct. Mol. Biol. **3**, 940-945.

62. Berezovskaya, Y., Armstrong, C. T., Boyle, A. L., Porrini, M., Woolfson, D. N. & Barran, P. E. (2011). Metal binding to a zinc-finger peptide: a comparison between solution and the gas phase. *Chem. Commun.* **47**, 412-414.

63. Mason, J. M. & Arndt, K. M. (2004). Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem* **5**, 170-176.

64. Pauling, L. & Corey, R. B. (1953). Compound helical configurations of polypeptide chains; structure of proteins of the O±-keratin type. *Nature (London, U. K.)* **171**, 59-61.

65. Crick, F. (1953). The packing of [alpha]-helices: simple coiled-coils. *Acta Crystallogr.* **6**, 689-697.

66. Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M. & Kim, P. S. (1995). Predicting coiled coils by use of pairwise residue correlations. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 8259-8263.

67. Gribbon, C., Channon, K. J., Zhang, W. J., Banwell, E. F., Bromley, E. H. C., Chaudhuri, J. B., Oreffo, R. O. C. & Woolfson, D. N. (2008). MagicWand: A single, designed peptide that assembles to stable, ordered alpha-helical fibers. *Biochemistry* **47**, 10365-10371.

68. Bromley, E. H. C., Channon, K., Moutevelis, E. & Woolfson, D. N. (2008). Peptide and protein building blocks for synthetic biology: From programming biomolecules to self-organized biomolecular systems. *ACS Chem. Biol.* **3**, 38-50.

69. O'Shea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide Velcro : Design of a heterodimeric coiled coil. *Curr. Biol.* **3**, 658-667.

70. Graddis, T. J., Myszka, D. G. & Chaiken, I. M. (1993). Controlled formation of model homo- and heterodimer coiled coil polypeptides. *Biochemistry* **32**, 12664-12671.

71. Bromley, E. H. C., Channon, K. J., King, P. J. S., Mahmoud, Z. N., Banwell, E. F., Butler, M. F., Crump, M. P., Dafforn, T. R., Hicks, M. R., Hirst, J. D., Rodger, A. & Woolfson, D. N. (2010). Assembly Pathway of a Designed [alpha]-Helical Protein Fiber. *Biophys. J.* **98**, 1668-1676.

72. Grigoryan, G. & Keating, A. E. (2008). Structural specificity in coiled-coil interactions. *Curr. Opin. Struct. Biol.* **18**, 477-483.

73. Moutevelis, E. & Woolfson, D. N. (2009). A Periodic Table of Coiled-Coil Protein Structures. *J. Mol. Biol.* **385**, 726-732.

74. Pandya, M. J., Spooner, G. M., Sunde, M., Thorpe, J. R., Rodger, A. & Woolfson, D. N. (2000). Sticky-End Assembly of a Designed Peptide Fiber Provides Insight into Protein Fibrillogenesis. *Biochemistry* **39**, 8728-8734.

75. Ryadnov, M. G. & Woolfson, D. N. (2004). Fiber Recruiting Peptides: Noncovalent Decoration of an Engineered Protein Scaffold. *J. Am. Chem. Soc.* **126**, 7454-7455.

76. Woolfson, D. N. & Ryadnov, M. G. (2006). Peptide-based fibrous biomaterials: some things old, new and borrowed. *Curr. Opin. Chem. Biol.* **10**, 559-567.

77. Banwell, E. F., Abelardo, E. S., Adams, D. J., Birchall, M. A., Corrigan, A., Donald, A. M., Kirkland, M., Serpell, L. C., Butler, M. F. & Woolfson, D. N. (2009). Rational design and application of responsive [alpha]-helical peptide hydrogels. *Nat. Mater.* **8**, 596-600.

78. Hall, C. K. (2008). Thermodynamic and kinetic origins of Alzheimer's and related diseases: A chemical engineer's perspective. *AlChE J.* **54**, 1956-1962.

79. Krylov, D., Barchi, J. & Vinson, C. (1998). Inter-helical interactions in the leucine zipper coiled coil dimer: ph and salt dependence of coupling energy between charged amino acids. *J. Mol. Biol.* **279**, 959-972.

80. Bromley, E. H. C., Sessions, R. B., Thomson, A. R. & Woolfson, D. N. (2009). Designed α-Helical Tectons for Constructing Multicomponent Synthetic Biological Systems. *J. Am. Chem. Soc.* **131**, 928-930.

81. Acharya, A., Rishi, V. & Vinson, C. (2006). Stability of 100 Homo and Heterotypic Coiled-Coil a-a' Pairs for Ten Amino Acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry* **45**, 11324-11332.

82. Mason, J. M., Schmitz, M. A., Mueller, K. M. & Arndt, K. M. (2006). Semirational design of Jun-Fos coiled coils with increased affinity: universal implications for leucine zipper prediction and design. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8989-8994.

83. Woolfson, D. N. & Alber, T. (1995). Predicting Oligomerization States of Coiled Coils. *Protein Sci.* **4**, 1596-1607.

84. Gernert, K. M., Surles, M. C., Labean, T. H., Richardson, J. S. & Richardson, D. C. (1995). The Alacoil: a very tight, antiparallel coiled-coil of helices. *Protein Sci.* **4**, 2252-60.

85. Gerstein, M. & Echols, N. (2004). Exploring the range of protein flexibility, from a structural proteomics perspective. *Curr. Opin. Chem. Biol.* **8**, 14-19.

86. Carrell, R. W. & Gooptu, B. (1998). Conformational changes and disease - serpins, prions and Alzheimer's. *Curr. Opin. Struct. Biol.* **8**, 799-809.

87. Carrell, R. W., Huntington, J. A., Mushunje, A. & Zhou, A. (2001). The conformational basis of thrombosis. *Thromb. Haemostasis* **86**, 14-22.

88. Beckett, D. (2004). Functional Switches in Transcription Regulation; Molecular Mimicry and Plasticity in Protein-Protein Interactions. *Biochemistry* **43**, 7983-7991.

89. Goh, C.-S., Milburn, D. & Gerstein, M. (2004). Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.* **14**, 104-109.

90. Schueler-Furman, O., Wang, C., Bradley, P., Misura, K. & Baker, D. (2005). Progress in Modeling of Protein Structures and Interactions. *Science (Washington, DC, U. S.)* **310**, 638-642.

91. Xia, Y. & Levitt, M. (2004). Simulating protein evolution in sequence and structure space. *Curr. Opin. Struct. Biol.* **14**, 202-207.

92. Mutter, M., Gassmann, R., Buttkus, U. & Altmann, K. H. (1991). Switch-peptides: pH induced transition of α-helix to β-folded sheet structure in bisamphiphilic oligopeptides. *Angew. Chem.* **103**, 1504-6 (See also Angew. Chem., Int. Ed. Engl., 1991, (30)11, 1514-16).

93. Pagel, K. & Koksch, B. (2008). Following polypeptide folding and assembly with conformational switches. *Curr. Opin. Chem. Biol.* **12**, 730-739.

94. Alexander, P. A., He, Y., Chen, Y., Orban, J. & Bryan, P. N. (2009). A minimal sequence code for switching protein structure and function. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21149-21154.

95. Mimna, R., Camus, M. S., Schmid, A., Tuchscherer, G., Lashuel, H. & Mutter, M. (2007). Disruption of Amyloid-Derived Peptide Assemblies through the Controlled Induction of a β-Sheet to α-Helix Transformation: Application of the Switch Concept. *Angew. Chem. Int. Ed.* **46**, 2681-2684.

96. Shi, Y., Beger, R. D. & Berg, J. M. (1993). Metal binding properties of single amino acid deletion mutants of zinc finger peptides: studies using cobalt(II) as a spectroscopic probe. *Biophys. J.* **64**, 749-753.

97. Ghosh, D. & Pecoraro, V. L. (2004). Understanding Metalloprotein Folding Using a de Novo Design Strategy. *Inorg. Chem.* **43**, 7902-7915.

98. Ghosh, D. & Pecoraro, V. L. (2005). Probing metal-protein interactions using a de novo design approach. *Curr. Opin. Chem. Biol.* **9**, 97-103.

99. Nitz, M., Sherawat, M., Franz, K. J., Peisach, E., Allen, K. N. & Imperiali, B. (2004). Structural origin of the high affinity of a chemically evolved lanthanide-binding peptide. *Angew. Chem., Int. Ed.* **43**, 3682-3685.

100. Tanaka, T., Mizuno, T., Fukui, S., Hiroaki, H., Oku, J., Kanaori, K., Tajima, K. & Shirakawa, M. (2004). Two-Metal Ion, Ni(II) and Cu(II), Binding α-Helical Coiled Coil Peptide. *J. Am. Chem. Soc.* **126**, 14023-14028.

101. Cerasoli, E., Sharpe, B. K. & Woolfson, D. N. (2005). ZiCo: A peptide designed to switch folded state upon binding zinc. *J. Am. Chem. Soc.* **127**, 15008-15009.

102. Kunz, W., Henle, J. & Ninham, B. W. (2004). 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. *Curr. Opin. Colloid Interface Sci.* **9**, 19-37.

103. Smith, J. D., Saykally, R. J. & Geissler, P. L. (2007). The Effects of Dissolved Halide Anions on Hydrogen Bonding in Liquid Water. *J. Am. Chem. Soc.* **129**, 13847-13856.

104. Zhang, Y. & Cremer, P. S. (2010). Chemistry of hofmeister anions and osmolytes. *Annu. Rev. Phys. Chem.* **61**, 63-83.

105. Ries-Kautt, M. M. & Ducruix, A. F. (1989). Relative effectiveness of various ions on the solubility and crystal growth of lysozyme. *J. Biol. Chem.* **264**, 745-8.

106. Bostroem, M., Tavares, F. W., Finet, S., Skouri-Panet, F., Tardieu, A. & Ninham, B. W. (2005). Why forces between proteins follow different Hofmeister series for pH above and below pI. *Biophys. Chem.* **117**, 217-224.

107. Zhou, P., Tian, F., Zou, J., Ren, Y., Liu, X. & Shang, Z. (2010). Do Halide Motifs Stabilize Protein Architecture? *J. Phys. Chem. B* **114**, 15673-15686.

108. Collins, K. D., Neilson, G. W. & Enderby, J. E. (2007). Ions in water: Characterizing the forces that control chemical processes and biological structure. Biophys. Chem. **128**, 95-104.

109. Tobias, D. J. & Hemminger, J. C. (2008). Getting specific about specific ion effects. *Science (Washington, DC, U. S.)* **319**, 1197-1198.

110. Koelsch, P., Viswanath, P., Motschmann, H., Shapovalov, V. L., Brezesinski, G., Moehwald, H., Horinek, D., Netz, R. R., Giewekemeyer, K., Salditt, T., Schollmeyer, H., Von, K. R., Daillant, J. & Guenoun, P. (2007). Specific ion effects in physicochemical and biological systems: Simulations, theory and experiments. *Colloids Surf., A* **303**, 110-136.

111. Chen, X., Yang, T., Kataoka, S. & Cremer, P. S. (2007). Specific ion effects on interfacial water structure near macromolecules. *J. Am. Chem. Soc.* **129**, 12272-12279.

112. Vrbka, L., Jungwirth, P., Bauduin, P., Touraud, D. & Kunz, W. (2006). Specific Ion Effects at Protein Surfaces: A Molecular Dynamics Study of Bovine Pancreatic Trypsin Inhibitor and Horseradish Peroxidase in Selected Salt Solutions. *J. Phys. Chem. B* **110**, 7036-7043.

113. Lund, M., Vacha, R. & Jungwirth, P. (2008). Specific ion binding to macromolecules: Effects of hydrophobicity and ion pairing. *Langmuir* **24**, 3387-3391.

114. Lund, M., Jungwirth, P. & Woodward, C. E. (2008). Ion Specific Protein Assembly and Hydrophobic Surface Forces. *Phys. Rev. Lett.* **100**, 258105/1-258105/4.

115. Jungwirth, P. & Winter, B. (2008). Ions at aqueous interfaces: From water surface to hydrated proteins. *Annu. Rev. Phys. Chem.* **59**, 343-366.

116. Cho, Y., Zhang, Y., Christensen, T., Sagle, L. B., Chilkoti, A. & Cremer, P. S. (2008). Effects of Hofmeister Anions on the Phase Transition Temperature of Elastin-like Polypeptides. *J. Phys. Chem. B* **112**, 13765-13771.

117. Canfield, R. E. (1963). The Amino Acid Sequence of Egg White Lysozyme. *J. Biol. Chem.* **238**, 2698-2707.

118. Blake, C. C., Koenig, D. F., Mair, G. A., North, A. C., Phillips, D. C. & Sarma, V. R. (1965). Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature* **206**, 757-761.

119. Lund, M. & Jungwirth, P. (2008). Patchy proteins, anions and the Hofmeister series. *J. Phys. Condens. Matter* **20**, 494218.

120. Zhang, Y. & Cremer, P. S. (2009). The inverse and direct Hofmeister series for lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15249-15253.

121. Gokarn, Y. R., Fesinmeyer, R. M., Saluja, A., Razinkov, V., Chase, S. F., Laue, T. M. & Brems, D. N. (2011). Effective charge measurements reveal selective and preferential accumulation of anions, but not cations, at the protein surface in dilute salt solutions. *Protein Sci.* **20**, 580-587.

122. Kulkarni, S. K., Ashcroft, A. E., Carey, M., Masselos, D., Robinson, C. V. & Radford, S. E. (1999). A near-native state on the slow refolding pathway of hen lysozyme. *Protein Sci.* **8**, 35-44.

123. Merenbloom, S., Flick, T., Daly, M. & Williams, E. (2011). Effects of Select Anions from the Hofmeister Series on the Gas-Phase Conformations of Protein Ions Measured with Traveling-Wave Ion Mobility Spectrometry/Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **22**, 1978-1990.

124. Li, Y., Hsieh, Y., Henion, J. & Ganem, B. (1993). Studies on heme binding in myoglobin, hemoglobin, and cytochrome c by ion spray mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **4**, 631-637.

125. Margoliash, E. (1962). Amino acid sequence of chymotryptic peptides from horse heart cytochrome c. *J. Biol. Chem.* **237**, 2161-2174.

126. Pelletier, H. & Kraut, J. (1992). Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c. *Science (Washington, D. C., 1883-)* **258**, 1748-55.

127. Baglioni, P., Fratini, E., Lonetti, B. & Chen, S. H. (2004). Structural arrest in concentrated cytochrome C solutions: the effect of pH and salts. *J. Phys. Condens. Matter* **16**, S5003-S5022.

128. Baglioni, P., Fratini, E., Lonetti, B. & Chen, S. H. (2004). Gelation in cytochrome C concentrated solutions near the isoelectric point: the anion role. *Curr. Opin. Colloid Interface Sci.* **9**, 38-42.

129. Battistuzzi, G., Borsari, M. & Sola, M. (1997). Anion binding to cytochrome c2: implications on protein-ion interactions in class I cytochromes c. *Arch. Biochem. Biophys.* **339**, 283-290.

130. Han, L., Hyung, S.-J., Mayers, J. J. S. & Ruotolo, B. T. (2011). Bound Anions Differentially Stabilize Multiprotein Complexes in the Absence of Bulk Solvent. *J. Am. Chem. Soc.* **133**, 11358-11367.

131. Kassell, B., Radicevic, M., Ansfield, M. J. & Laskowski, M., Sr. (1965). Basic trypsin inhibitor of bovine pancreas. IV. Linear sequence of the 58 amino acids. *Biochem. Biophys. Res. Commun.* **18**, 255-8.

132. Deisenhofer, J. & Steigemann, W. (1975). Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1.5 Ang. resolution. *Acta Crystallogr., Sect. B* **B31**, 238-50.

133. Ascenzi, P., Bocedi, A., Bolognesi, M., Spallarossa, A., Coletta, M., De, C. R. & Menegatti, E. (2003). The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): A milestone protein. *Curr. Protein Pept. Sci.* **4**, 231-251.

134. Kuroda, Y. & Kim, P. S. (2000). Folding of bovine pancreatic trypsin inhibitor (BPTI) variants in which almost half the residues are alanine. *J. Mol. Biol.* **298**, 493-501.

135. Shaw, D. E., Maragakis, P., Lindorff-Larsen, K., Piana, S., Dror, R. O., Eastwood, M. P., Bank, J. A., Jumper, J. M., Salmon, J. K., Shan, Y. & Wriggers, W. (2010). Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **330**, 341-346.

136. Shelimov, K. B., Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. (1997). Protein Structure in Vacuo: Gas-Phase Conformations of BPTI and Cytochrome c. *J. Am. Chem. Soc.* **119**, 2240-2248.

137. Creighton, T. E. (1980). Role of the environment in the refolding of reduced pancreatic trypsin inhibitor. *J. Mol. Biol.* **144**, 521-50.

138. Veesler, S., Lafont, S., Marcq, S., Astier, J. P. & Boistelle, R. (1996). Prenucleation, crystal growth and polymorphism of some proteins. *J. Cryst. Growth* **168**, 124-129.

139. Neet, K. E. & Lee, J. C. (2002). Biophysical characterization of proteins in the post-genomic era of proteomics. *Mol. Cell. Proteomics* **1**, 415-420.

140. Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D. & Alice, M. B. (1968). Molecular Beams of Macroions. *J. Chem. Phys.* **49**, 2240-2249.

141. Yamashita, M. & Fenn, J. B. (2002). Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **88**, 4451-4459.

142. Loo, J. A. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16**, 1-23.

143. Sterling, H., Batchelor, J., Wemmer, D. & Williams, E. (2010). Effects of buffer loading for electrospray ionization mass spectrometry of a noncovalent protein complex that requires high concentrations of essential salts. *J. Am. Soc. Mass. Spectrom.* **21**, 1045-1049.

144. Liu, C., Wu, Q., Harms, A. C. & Smith, R. D. (1996). On-line microdialysis sample cleanup for electrospray ionization mass spectrometry of nucleic acid samples. *Anal. Chem.* **68**, 3295-9.

145. Chen, Y., Mori, M., Pastusek, A. C., Schug, K. A. & Dasgupta, P. K. (2010). On-Line Electrodialytic Salt Removal in Electrospray Ionization Mass Spectrometry of Proteins. *Anal. Chem.* **83**, 1015-1021.

146. Rostom, A. A. & Robinson, C. V. (1999). Detection of the Intact GroEL Chaperonin Assembly by Mass Spectrometry. *J. Am. Chem. Soc.* **121**, 4718-4719.

147. Briand, G., Perrier, V., Kouach, M., Takahashi, M., Gilles, A.-M. & Bârzu, O. (1997). Characterization of Metal and Nucleotide Liganded Forms of Adenylate Kinase by Electrospray Ionization Mass Spectrometry. *Arch. Biochem. Biophys.* **339**, 291-297.

148. Fundel, S. M., Pountney, D. L., Bogumil, R., Gehrig, P. M., Hasler, D. W., Faller, P. & Vasak, M. (1996). Isolation and characterization of a novel monomeric zinc- and heme-containing protein from bovine brain. *FEBS Lett.* **395**, 33-38.

149. Wyttenbach, T., Grabenauer, M., Thalassinos, K., Scrivens, J. H. & Bowers, M. T. (2009). The Effect of Calcium Ions and Peptide Ligands on the Relative Stabilities of the Calmodulin Dumbbell and Compact Structures. *J. Phys. Chem. B* **114**, 437-447.

150. Wyttenbach, T., Batka, J. J., Jr., Gidden, J. & Bowers, M. T. (1999). Host/guest conformations of biological systems: valinomycin/alkali ions. *Int. J. Mass Spectrom.* **193**, 143-152.

151. Stephenson, J. L. & McLuckey, S. A. (1997). Counting Basic Sites in Oligopeptides via Gas-Phase Ion Chemistry. *Anal. Chem.* **69**, 281-285.

152. Lei, Q. P., Cui, X., Kurtz, D. M., Amster, I. J., Chernushevich, I. V. & Standing, K. G. (1998). Electrospray Mass Spectrometry Studies of Non-Heme Iron-Containing Proteins. *Anal. Chem.* **70**, 1838-1846.

153. Lim, N. C., Freake, H. C. & Brueckner, C. (2005). Illuminating zinc in biological systems. *Chem. Eur. J.* **11**, 38-49.

154. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L. & Berg, J. M. (1991). A consensus zinc finger peptide: design, high-affinity metal binding, a pH-dependent structure, and a His to Cys sequence variant. *J. Am. Chem. Soc.* **113**, 4518-4523.

155. Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T. & Berg, J. M. (1992). Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4796-800.

156. Thomson, J. J. (1913). *Rays of Positive Electricity and their Application to Chemical Analyses*, Longmans, Green and Company: London, U.K., 1913.

157. de Hoffmann, E. & Stroobant, V. (2002). *Mass Spectrometery: Principles and Applications.* 2nd Edition edit, John Wiley & Sons, Ltd.

158. Conrad, R. (1934). Decomposition of hexane, cyclohexane and benzene in the positive ray tube. *Trans. Faraday Soc.* **30**, 0215-0220.

159. Barber, M., Bordoli, R. S., Sedgwick, R. D. & Tyler, A. N. (1981). Fast atom bombardment of solids as an ion source in mass spectrometry. *Nature (London)* **293**, 270-5.

160. Desiderio, D. M. & Katakuse, I. (1984). Fast atom bombardment mass spectrometry of insulin, insulin A-chain, insulin B-chain, and glucagon. *Biomed. Mass Spectrom.* **11**, 55-9.

161. Fenn, J. B. (2002). Electrospray Ionization Mass Spectrometry: How It All Began. *J. Biomol. Techniques* **13**, 101-118.

162. Karas, M. & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**, 2299-2301.

163. Tanaka, K., Ido, Y., Akita, S., Yoshida, Y. & Yoshida, T. (1987). *Proc. Second Japan-China Symp. Mass Spectrom.*

164. Karas, M., Bachmann, D. & Hillenkamp, F. (1985). Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem.* **57**, 2935-2939.

165. Fenn, J. B. (2002). Electrospray Wings for Molecular Elephants. In *Les Prix Nobel*, pp. 154-184.

166. Gomez, A. & Tang, K. (1994). Charge and fission of droplets in electrostatic sprays. *Phys. Fluids* **6**, 404-414.

167. Ikonomou, M. G., Blades, A. T. & Kebarle, P. (1991). Electrospray-ion spray: a comparison of mechanisms and performance. *Anal. Chem.* **63**, 1989-98.

168. Wilm, M. S. & Mann, M. (1994). Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Processes* **136**, 167-80.

169. Smith, R. D., Shen, Y. & Tang, K. (2004). Ultrasensitive and Quantitative Analyses from Combined Separations–Mass

Spectrometry for the Characterization of Proteomes. *Acc. Chem. Res.* **37**, 269-278.

170. Wilm, M. & Mann, M. (1996). Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **68**, 1-8.

171. Benkestock, K., Sundqvist, G., Edlund, P.-O. & Roeraade, J. (2004). Influence of droplet size, capillary-cone distance and selected instrumental parameters for the analysis of noncovalent protein-ligand complexes by nano-electrospray ionization mass spectrometry. *J. Mass Spectrom.* **39**, 1059-1067.

172. Gaskell, S. J. (1985). Mass spectrometry in medical research. *Clin. Phys. Physiol. Meas.* **6**, 1.

173. Khalsa-Moyers, G. & McDonald, W. H. (2006). Developments in mass spectrometry for the analysis of complex protein mixtures. *Brief. Funct. Genomics Proteomics* **5**, 98-111.

174. Siuzdak, G. (2006). *The Expanding Role of Mass Spectrometry in Biotechnology*. 2nd edn. edit, MCC Press, San Diego.

175. Paul, W. & Steinwedel, H. (1953). A new mass spectrometer without magnetic field. *Z. Naturforsch.* **8a**, 448-50.

176. Finnigan, R. E. (1994). Quadrupole mass spectrometers. *Anal. Chem.* **66**, 969A-975A.

177. Kicman, A. T., Parkin, M. C. & Iles, R. K. (2007). An introduction to mass spectrometry based proteomics—Detection and characterization of gonadotropins and related molecules. *Mol. Cell. Endocrinol.* **260–262**, 212-227.

178. Stephens, W. E. (1946). A pulsed mass spectrometer with time dispersion. *Phys. Rev.* **69**, 691.

179. Wiley, W. C. & McLaren, I. H. (1955). Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* **26**, 1150-7.

180. Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. & Zagulin, V. A. (1973). Mass reflectron. New nonmagnetic time-of-flight high-resolution mass spectrometer. *Zh. Eksp. Teor. Fiz.* **64**, 82-9.

181. Soares, R., Franco, C., Pires, E., Ventosa, M., Palhinhas, R., Koci, K., Martinho de Almeida, A. & Varela Coelho, A. (2012). Mass spectrometry and animal science: Protein identification strategies and particularities of farm animal species. *J. Proteomics* **In press**.

182. Iams, H. & Salzberg, B. (1935). The Secondary Emission Phototube. *Proc. Inst. Radio Eng.* **23**, 55-64.

183. Wiza, J. L. (1979). Microchannel plate detectors. *Nucl. Instrum. Methods* **162**, 587-601.

184. Shukla, A. K. & Futrell, J. H. (2000). Tandem mass spectrometry: dissociation of ions by collisional activation. *J. Mass Spectrom.* **35**, 1069-1090.

185. Sleno, L. & Volmer, D. A. (2004). Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **39**, 1091-1112.

186. Haddon, W. F. & McLafferty, F. W. (1968). Metastable ion characteristics. VII. Collision-induced metastables. *J. Am. Chem. Soc.* **90**, 4745-6.

187. Jennings, K. R. (1968). Collision-induced decompositions of aromatic molecular ions. *Int. J. Mass Spectrom. Ion Phys.* **1**, 227-35.

188. Cooks, R. G., Beynon, J. H., Caprioli, R. M. & Lester, G. R. (1973). *Metastable Ions*, Elsevier.

189. Biemann, K., Cone, C., Webster, B. R. & Arsenault, G. P. (1966). Determination of the amino acid sequence in oligopeptides by computer interpretation of their high-resolution mass spectra. *J. Am. Chem. Soc.* **88**, 5598-606.

190. Smith, D. L. & Zhang, Z. (1994). Probing noncovalent structural features of proteins by mass spectrometry. *Mass Spectrom. Rev.* **13**, 411-29.

191. Roepstorff, P. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11**, 601.

192. Benesch, J. L. P. (2009). Collisional Activation of Protein Complexes: Picking Up the Pieces. *J. Am. Soc. Mass Spectrom.* **20**, 341-348.

193. Guevremont, R., Siu, K. W. M., Wang, J. & Ding, L. (1997). Combined Ion Mobility/Time-of-Flight Mass Spectrometry Study of Electrospray-Generated Ions. *Anal. Chem.* **69**, 3959-3965.

194. Hoaglund, C. S., Valentine, S. J., Sporleder, C. R., Reilly, J. P. & Clemmer, D. E. (1998). Three-Dimensional Ion Mobility/TOFMS Analysis of Electrosprayed Biomolecules. *Anal. Chem.* **70**, 2236-2242.

195. Langevin, P. (1903). Ionization of Gas. [machine translation]. *Ann. Chim. Phys.* **28**, 289-384.

196. Langevin, P. (1905). A fundamental formula of kinetic theory. *Ann. Chim. Phys.* **5**, 245-288.

197. Cravath, A. M. (1929). The rate of formation of negative ions by electron attachment. *Phys. Rev.* **33**, 605-13.

198. van de Graaff, R. J. (1929). Mobility of ions in gases. *Nature (London, U. K.)* **124**, 10-11.

199. Eiceman, G. A. & Karpas, Z. (2005). *Ion Mobility Mass Spectrometry*. 2nd Edition edit, CRC Press.

200. Clemmer, D. E. & Jarrold, M. F. (1997). Ion Mobility Measurements and their Applications to Clusters and Biomolecules. *J. Mass Spectrom.* **32**, 577-592.

201. Fernandez-Lima, F. A., Wei, H., Gao, Y. Q. & Russell, D. H. (2009). On the Structure Elucidation Using Ion Mobility Spectrometry and Molecular Dynamics. *J. Phys. Chem. A* **113**, 8221-8234.

202. Mason, E. A. & McDaniel, E. W. (1988). *Transport Properties of Ions in Gases*, Wiley: New York.

203. Wyttenbach, T. & Bowers, M. (2003). Gas-Phase Conformations: The Ion Mobility/Ion Chromatography Method. In *Modern Mass Spectrometry*, pp. 207-232.

204. Creaser, C. S., Griffiths, J. R., Bramwell, C. J., Noreen, S., Hill, C. A. & Thomas, C. L. P. (2004). Ion mobility spectrometry: a review. Part 1. Structural analysis by mobility measurement. *Analyst* **129**, 984-994.

205. Bohrer, B. C., Merenbloom, S. I., Koeniger, S. L., Hilderbrand, A. E. & Clemmer, D. E. (2008). Biomolecule Analysis by Ion Mobility Spectrometry. *Annu. Rev. Anal. Chem.* **1**, 293-327.

206. Verbeck, G. F., Ruotolo, B.T., Sawyer, H.A., Gillig, K.J., and Russell, D.H. (2002). A Fundamental Introduction to Ion Mobility Mass Spectrometry Applied to the Analysis of Biomolecules. *J. Biomol. Techniques* **13**, 56-61.

207. Kaddis, C. S. & Loo, J. A. (2007). Native protein MS and Ion mobility: Large flying proteins with ESI. *Anal. Chem.* **79**, 1778-1784.

208. Dugourd, P., Hudgins, R. R., Clemmer, D. E. & Jarrold, M. F. (1997). High-resolution ion mobility measurements. *Rev. Sci. Instrum.* **68**, 1122-1129.

209. Purves, R. W., Guevremont, R., Day, S., Pipich, C. W. & Matyjaszczyk, M. S. (1998). Mass spectrometric characterization of a high-field asymmetric waveform ion mobility spectrometer. *Rev. Sci. Instrum.* **69**, 4094-4105.

210. Shvartsburg, A. A. & Smith, R. D. (2007). Scaling of the Resolving Power and Sensitivity for Planar FAIMS and Mobility-Based

Discrimination in Flow- and Field-Driven Analyzers. *J. Am. Soc. Mass Spectrom.* **18**, 1672-1681.

211. Guevremont, R. (2004). High-field asymmetric waveform ion mobility spectrometry: A new tool for mass spectrometry. *J. Chromatogr. A* **1058**, 3-19.

212. Purves, R. W. & Guevremont, R. (1999). Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry - Mass Spectrometry. *Anal. Chem.* **71**, 2346-2357.

213. Guevremont, R., Barnett, D. A., Purves, R. W. & Vandermey, J. (2000). Analysis of a Tryptic Digest of Pig Hemoglobin Using ESI-FAIMS-MS. *Anal. Chem.* **72**, 4577-4584.

214. Barnett, D. A., Belford, M., Dunyach, J.-J. & Purves, R. W. (2007). Characterization of a Temperature-Controlled FAIMS System. *J. Am. Soc. Mass Spectrom.* **18**, 1653-1663.

215. Shvartsburg, A. A., Tang, K., Smith, R. D., Holden, M., Rush, M., Thompson, A. & Toutoungi, D. (2009). Ultrafast Differential Ion Mobility Spectrometry at Extreme Electric Fields Coupled to Mass Spectrometry. *Anal. Chem. (Washington, DC, U. S.)* **81**, 8048-8053.

216. Purves, R. W., Barnett, D. A. & Guevremont, R. (2000). Separation of protein conformers using electrospray-high field asymmetric waveform ion mobility spectrometry-mass spectrometry. *Int. J. Mass Spectrom.* **197**, 163-177.

217. Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinos, K., Bateman, R. H., Bowers, M. T. & Scrivens, J. H. (2007). An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom.* **261**, 1-12.

218. Shvartsburg, A. A. & Smith, R. D. (2008). Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem. (Washington, DC, U. S.)* **80**, 9689-9699.

219. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. & Robinson, C. V. (2008). Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* **3**, 1139-1152.

220. Merenbloom, S., Flick, T. & Williams, E. (2011). How Hot are Your Ions in TWAVE Ion Mobility Spectrometry? *J. Am. Soc. Mass. Spectrom.*, 1-10.

221. Chowdhury, S. K., Katta, V. & Chait, B. T. (1990). Probing conformational changes in proteins by mass spectrometry. *J. Am. Chem. Soc.* **112**, 9012-13.

222. Przybylski, M. & Glocker, M. O. (1996). Electrospray mass spectrometry of biomacromolecular complexes with noncovalent interactions - new analytical perspectives for supramolecular chemistry and molecular recognition processes. *Angew. Chem., Int. Ed. Engl.* **35**, 806-826.

223. Afonso, C., Hathout, Y. & Fenselau, C. (2002). Qualitative characterization of biomolecular zinc complexes by collisionally induced dissociation. *J. Mass Spectrom.* **37**, 755-759.

224. Fabris, D. (2005). Mass spectrometric approaches for the investigation of dynamic processes in condensed phase. *Mass Spectrom. Rev.* **24**, 30-54.

225. Want, E. J., Cravatt, B. F. & Siuzdak, G. (2005). The expanding role of mass spectrometry in metabolite profiling and characterization. *ChemBioChem* **6**, 1941-1951.

226. Konermann, L., Pan, J. & Wilson, D. J. (2005). Protein folding mechanisms studied by time-resolved electrospray mass spectrometry. *BioTechniques* **40**, 135,137,139,141.

227. Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H. & Robinson, C. V. (2005). Evidence for macromolecular protein rings in the absence of bulk water. *Science* **310**, 1658-1661.

228. Loo, J. A., Berhane, B., Kaddis, C. S., Wooding, K. M., Xie, Y., Kaufman, S. L. & Chernushevich, I. V. (2005). Electrospray Ionization Mass Spectrometry and Ion Mobility Analysis of the 20S Proteasome Complex. *J. Am. Soc. Mass. Spectrom.* **16**, 998-1008.

229. Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. (1995). Naked Protein Conformations - Cytochrome-C in the Gas-Phase. *J. Am. Chem. Soc.* **117**, 10141-10142.

230. Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. (2007). Protein complexes in the gas phase: Technology for structural genomics and proteomics. *Chem. Rev.* **107**, 3544-3567.

231. Sharon, M. & Robinson, C. V. (2007). The role of mass Spectrometry in structure elucidation of dynamic protein complexes. *Annu. Rev. Biochem.* **76**, 167-193.

232. Valentine, S. J., Counterman, A. E. & Clemmer, D. E. (1999). A database of 660 peptide ion cross sections: use of intrinsic size parameters for bona fide predictions of cross sections. *J. Am. Soc. Mass. Spectrom.* **10**, 1188-1211.

233. De Cecco, M., Seo, E. S., Clarke, D. J., McCullough, B. J., Taylor, K., Macmillan, D., Dorin, J. R., Campopiano, D. J. & Barran, P. E. (2010). Conformational Preferences of Linear OI-Defensins Are Revealed by Ion Mobility-Mass Spectrometry. *J. Phys. Chem. B* **114**, 2312-2318.

234. Fuerstenau, S. D., Benner, W. H., Thomas, J. J., Brugidou, C., Bothner, B. & Siuzdak, G. (2001). Mass spectrometry of an intact virus. *Angew. Chem., Int. Ed.* **40**, 541-544.

235. Wuethrich, K. (1990). Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* **265**, 22059-62.

236. Jehle, S., Falb, M., Kirkpatrick, J. P., Oschkinat, H., van, R. B.-J., Althoff, G. & Carlomagno, T. (2010). Intermolecular Protein-RNA Interactions Revealed by 2D 31P-15N Magic Angle Spinning Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **132**, 3842-3846.

237. Schmid, F.-X. (2001). Biological macromolecules: UV-visible spectrophotometry. In *Encyclopedia of Life Sciences*, Vol. 2, pp. 965-968. Copyright (C) 2012 American Chemical Society (ACS). All Rights Reserved. vols. Macmillan Publishers Ltd, Nature Publishing Group.

238. Kelly, S. M., Jess, T. J. & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochim. Biophys. Acta, Proteins Proteomics* **1751**, 119-139.

239. Younan, N. D., Klewpatinond, M., Davies, P., Ruban, A. V., Brown, D. R. & Viles, J. H. (2011). Copper(II)-Induced Secondary Structure Changes and Reduced Folding Stability of the Prion Protein. *J. Mol. Biol.* **410**, 369-382.

240. Ralston, G. (2004). *Introduction to Analytical Ultracentrifugation*, Department of Biochemistry, The University of Sydney, Sydney, Australia.

241. Petrova, T. & Podjarny, A. (2004). Protein crystallography at subatomic resolution. *Rep. Prog. Phys.* **67**, 1565-1605.

242. Ennifar, E. (2012). X-ray Crystallography as a tool for Mechanism-of-action Studies and Drug Discovery. *Curr. Pharm. Biotechnol.*

243. Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C. & Jarrold, M. F. (1996). Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. *J. Phys. Chem.* **100**, 16082-16086.

244. Shvartsburg, A. A. & Jarrold, M. F. (1996). An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86-91.

245. von Helden, G., Hsu, M. T., Gotts, N. G., Kemper, P. R. & Bowers, M. T. (1993). Do small fullerenes exist only on the computer? Experimental results on C20 and C24 carbon cluster monopositive and mononegative ions. *Chem. Phys. Lett.* **204**, 15-22.

246. Tobias, D. J., Mertz, J. E. & Brooks III, C. L. (1991). Nanosecond Time Scale Folding Dynamics of a Pentapeptide in Water. *Biochemistry* **30**, 6054-6058.

247. Pain, R. H., Ed. (2000). Mechanisms of Protein Folding. 2nd Edition edit. Edited by Hames, B. D. & Glover, D. M.: Oxford University Press.

248. Bussi, G., Laio, A. & Parrinello, M. (2006). Equilibrium Free Energies from Nonequilibrium Metadynamics. *Phys. Rev. Lett.* **96**, 090601/1-090601/4.

249. Shea, J.-E. & Brooks III, C. L. (2001). From Folding Theories to Folding Proteins: a Review and Assessment of Simulation Studies of Protein Folding and Unfolding. *Annu. Rev. Phys. Chem.* **52**, 499-535.

Chapter 2

Experimental

2.1 Biological mass spectrometry

The mass spectrometry work presented in this thesis has been carried out using a hybrid quadrupole time of flight mass spectrometer equipped with nano-electrospray ionisation source – the Q-ToF-2 (Quadrupole Time of Flight) instrument (Waters, Manchester, UK). All data were acquired in positive ionisation mode. A schematic diagram of the instrument in shown in ***Figure 2-1***.

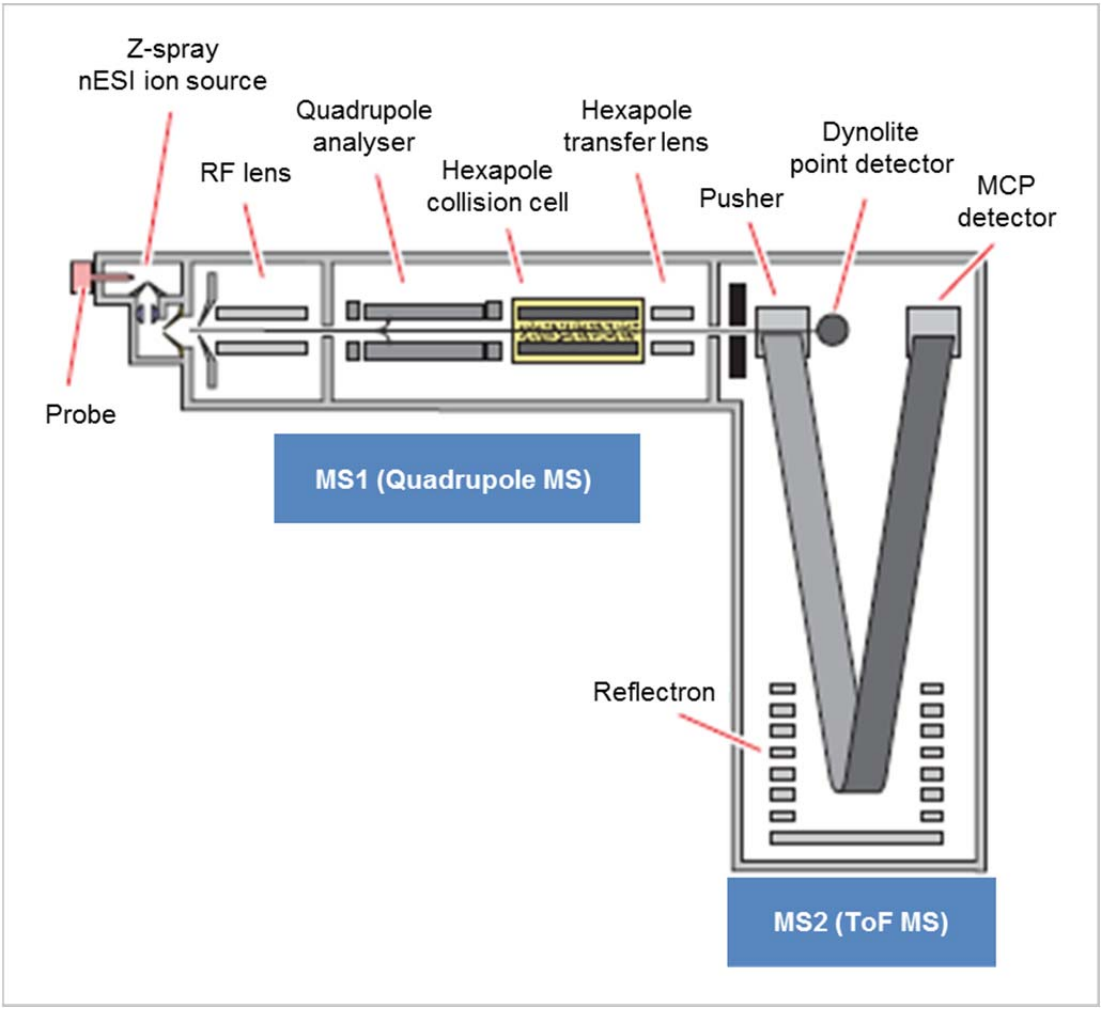


Figure 2-1 *Waters Q-ToF-2 mass spectrometer schematic (adapted from the manufacturer's user guide).The instrument consists of three differentially pumped regions: source, quadrupole analyser and ToF analyser, separated by small orifices.*

2.1.1 Nano-electrospray ionisation and sample introduction

Samples were ionised using nESI, where the proteinaceous species were charged and transferred from their solution into the gas phase. This was implemented using borosilicate capillaries with fine tips (1 – 30 μm) filled with sample solution into which 0.125 mm diameter platinum wire (Goodfellow Cambridge Ltd., Huntingdon, UK) was inserted to apply charge

to the analyte. Sample tips were fabricated in-house on a Flaming/Brown micropipette puller Model P-97 (Sutter Instrument Company, Novato, CA, USA) from thin-walled 0.9 mm i.d. capillaries (World precision Instruments, Sarasota, FL, USA). The voltage applied *via* a platinum wire was in the range of 0.7 – 1.5 kV and largely depended on the solvent composition and especially the spray tip geometry. The latter, although every care has been taken to ensure consistency of the capillary puller performance, was challenging to control accurately over the lifetime of the puller's heating filament.

2.1.2 The fate of ions in the mass spectrometer

The following describes the lifecycle of ions, from generation to detection, in the Q-ToF-2 mass spectrometer (see schematic in ***Figure 2-1***). The ions are generated in the source and enter the first vacuum region (~1.9 ± 16% mbar) of the mass spectrometer *via* a z-shaped trajectory afforded by mutually orthogonal positioning of the sprayer and the first two conical apertures – sample and extraction cones. Such source geometry improves sensitivity, as it allows for more efficient sample desolvation and removal of neutral species which are not sensitive to the voltages applied to both cones. The source block is typically heated to 80°C to improve desolvation. The ion beam is then focussed into the next vacuum region in the hexapole RF lens.

Upon entering the second vacuum region – the quadrupole analyser chamber – maintained at ~$2.7 \cdot 10^{-5}$ ± 13% mbar, the ions pass through a quadrupole. The operation of the quadrupole can be altered to accomplish either of the following experiments:

❖ MS analysis in the ToF. In this case, the quadrupole is used to transfer a pre-selected range of *m*/*z* values towards the ToF analyser. Here, the point detector is used to optimise the ion signal.

❖ MS/MS analysis of an ion of a selected *m*/*z*. Here, the quadrupole is tuned to transfer only ions with user-specified *m*/*z*. These ions can then be subjected to increased kinetic energy by applying voltages in the range of 4 – 200 V to be fragmented in the collision cell. The precursor ions of up to 4000 *m*/*z* can be selected by the quadrupole in this instrument.

In the latter case, the fragmentation occurs due to multiple collisions of the proteinaceous analyte ions with the neutral gas (argon) maintained at a pressure of 1 bar. This leads to fragmentation of the peptide bond along the protein's/peptide's backbone, generating primarily *b*- and *y*-type fragments (see section ***1.2.5*** for more details on fragment types). Any ions that survived and/or came to exist in the collision cell, whether the quadrupole was set to transfer (MS) or selection (MS/MS) mode, are transferred to the third pumping region *via* the RF hexapole lens.

This last – ToF – region is maintained at the highest vacuum of ~$4.6 \cdot 10^{-7}$ ± 26% mbar, to ensure minimal interference with the background ions and hence better analyte detection. In this region, the ions are analysed (or sorted) by measuring their time of flight, which is proportional to each ion's *m*/*z* ratio and inversely proportional to its velocity. On entry to this region, the ions are focussed and orthogonally accelerated by the pulsed pusher down the flight tube towards the reflectron ('ion mirror'). Here, they are turned around in the retarding field and directed to the micro-channel plates

set to 2000 V, where the ion signal is detected. Under the recommended instrument settings, the upper limit for this instrument is 32,000 *m*/*z*.

The detected ion signal is converted to the time of arrival for each incoming pulse in the time-to-digital converter (TDC) at a rate of 1 GHz. The resulting time domain is processed to yield the *m*/*z* domain by the MassLynx 4.1™ (Waters, Manchester, UK) built-in software and is presented as mass spectra.

2.1.3 Mass calibration

The mass spectrometer was calibrated daily using a 2 mg/mL sodium iodide dissolved in 70% isopropanol. The salt yields a series of cluster peaks corresponding to $Na_{n+1}I_n$ ions, where lowest value of *n* is 0, covering the whole *m*/*z* range accessible by the instrument. Typical peak resolution (FWHM) of the ToF signal was ~7800 on $(M + 6H)^{6+}$ ion from BPTI.

2.1.4 Typical Q-ToF-2 settings

Typical instrument settings used in the presented work are shown in ***Table 2-1***.

Parameter	Value	Comment
Capillary (kV)	1.00	
Cone	30	
Extractor	0	
RF Lens	0.40	
Source Temp (°C)	80	
LM Resolution	11	
HM Resolution	11	
Collision Energy	5	Raised stepwise for CID
Ion Energy	1.8	
Steering	–1.93	
Entrance	60	
Pre-filter	6.0	
Transport	2.0	5.0 for the study in ***Chapter 5***
Aperture2	9.0	
Acceleration	200	
Focus	0	
Tube Lens	100	
Offset1	–0.30	
Offset2	0.00	
Pusher	980	
TOF (kV)	9.10	
Reflectron	34.66	
Pusher Cycle Time (μs)	70 – 110	Adjusted according to *m*/*z* range
Multiplier	600	
MCP	2000	

Table 2-1 *Typical Q-ToF-2 settings. Values are shown in volts, unless stated otherwise.*

2.2 Ion mobility mass spectrometry: the MoQ-ToF

The IM-MS work presented in this thesis has been carried out on the 'MoQ-ToF' (acronym of 'Mobility Q-ToF'). This instrument is similar to the one described in section ***2.1.2***, and is based on the commercial Q-ToF mass spectrometer (Waters, Manchester, UK), modified in-house to enable ion

mobility separations. All IM-MS data were acquired in positive ionisation mode using nESI source as described in section ***2.1.1***.

2.2.1 Instrument description and operation

The main modification that turned the Q-ToF mass spectrometer into MoQ-ToF was the inclusion of the drift cell chamber containing the 5.1 cm drift cell, associated optics, gas feed-lines, *etc*. A schematic diagram of the MoQ-ToF is shown in ***Figure 2-2*** with the drift cell chamber highlighted in pale blue. The drift cell is filled with 3.2 – 4 Torr (4.3 – 5.3 mbar) buffer gas – helium (CP grade, 99.999% purity, BOC Specialty Gases Ltd, Guildford, UK) at a temperature of ~300 K (27°C) to enable mobility separations. This additional gas load in the middle of the vacuum region is dealt with by the 500 L/s Pfeiffer TMH520 turbomolecular pump (Pfeiffer Vacuum Ltd, Newport Pagnell, UK) backed by the Edwards two-stage E2M40 rotary vane pump (Edwards Vacuum, Crawley, UK); an additional mechanical booster pump is engaged when the drift cell is filled with gas. The resulting vacuum attained in the MoQ-ToF with the drift gas in the cell is as follows: source region: ~$1.9 \cdot 10^{-1}$ mbar; analyser region: ~$1.7 \cdot 10^{-3}$ mbar; ToF region: ~$1.4 \cdot 10^{-7}$ mbar.

The temperature of the drift cell can be varied between 80 and 700 K. Three different temperature settings have been used in this work: 'ambient' (~300 K), 'high' (~360 K) and 'low' (~260 K). Cell heating is implemented *via* tantalum wire (0.25 mm diameter, Goodfellow Cambridge Ltd., Huntingdon, UK) fed through ceramic rods (Kimball Physics Inc., Wilton, NH, USA) located in the copper block that houses the cell. A potential difference is

applied to the heaters causing the wire to resistively heat the cell copper block. Cell cooling is implemented *via* passing the chilled nitrogen gas through the channels inside the copper block. The outlets of the channels are connected to the ¼-inch o.d. stainless steel tubular cooling coil *via* Swagelok VCR vacuum-tight fittings (Swagelok Company, Kings Langley, UK). The gas is supplied from in-house reservoir through the pressure regulator kept at ~0.5 bar, is chilled in thermally insulated Dewar flask filled with liquid nitrogen before being supplied to the in-cell channels and exhausted into the atmosphere.

The MoQ-ToF can be operated either in the MS only mode or IM-MS mode, with the drift cell filled with helium in both cases. In the former case, the ion path is continuous and similar to the one described in section ***2.1.2***, whereas in the latter case a pulsed event is introduced in the drift cell to enable ion mobility separation of the analyte ions. MS mode is used to tune for ion signal, with the drift cell potential set at its highest value of 60 V (refer to section ***2.2.2*** for drift cell voltage range used). The drift cell pulse is synchronised with the ToF pulse to allow the drift time measurements, with one mobility separation event sampled 200 times (*i.e.* pushes) by the ToF.

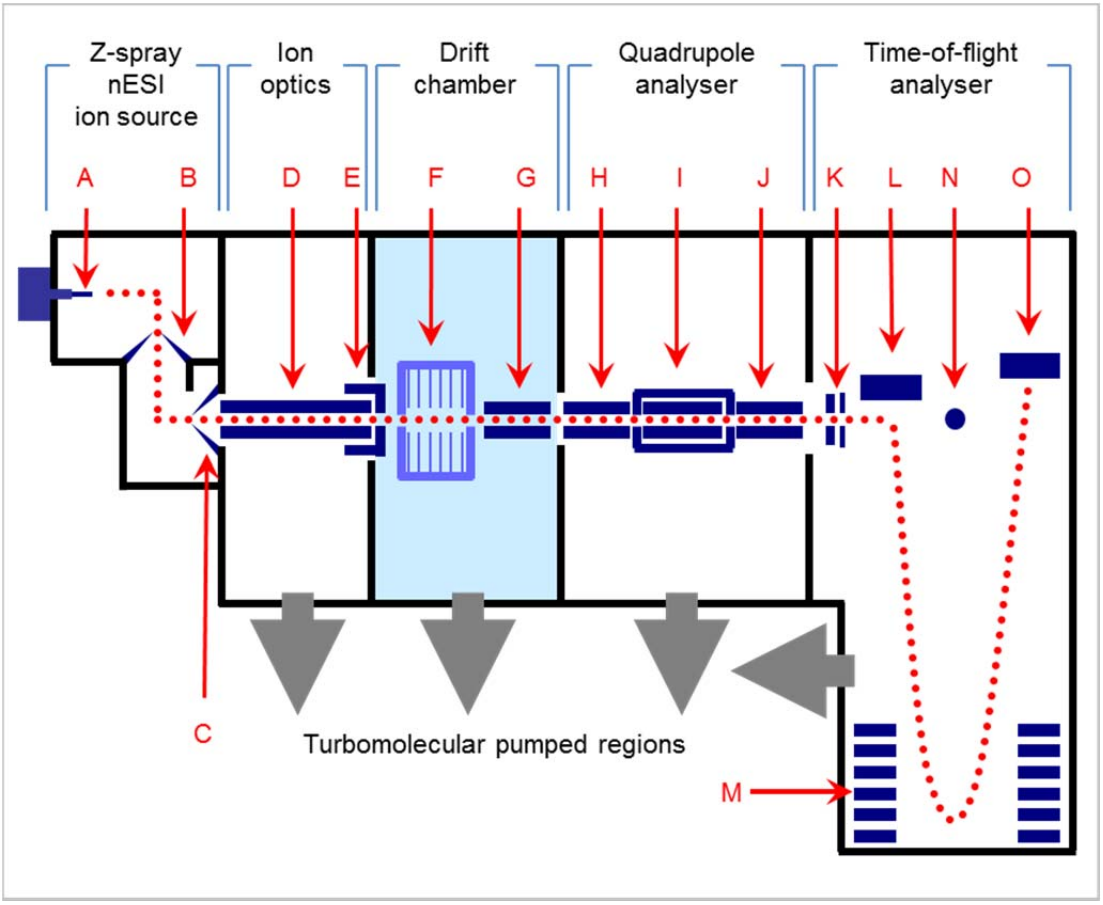


Figure 2-2 *The MoQ-ToF ion mobility mass spectrometer schematic with the ion path shown in red dotted line and the drift cell chamber highlighted in pale blue.* ***A*** *- electrospray capillary;* ***B*** *&* ***C*** *– sampling and extraction cones,* ***D*** *– pre-cell hexapole;* ***E*** *– ion gate ('top-hat' lens);* ***F*** *– drift cell containing 5 drift rings;* ***G*** *– post-cell hexapole;* ***H*** *– quadrupole;* ***I*** *– RF hexapole collision cell;* ***J*** *– transport RF hexapole;* ***K*** *– acceleration and focussing lenses;* ***L*** *– pusher lens stack;* ***M*** *– reflectron;* ***N*** *– point detector (for quadrupole analyser);* ***O*** *– MCP detector (for ToF analyser).*

Each mobility event starts with trapping of ions in the pre-cell hexapole (***Figure 2-2***) by raising the voltage on the 'top-hat' lens at the end of the hexapole. Then the trapping potential is lowered for a fixed time of 40 μs to allow the packet of ions enter the drift cell due to the potential difference between the hexapole and the cell entrance. This potential difference is termed 'injection energy' (IE) and is typically in the range of 35 – 40 V. Lower values of IE will cause insufficient ion transfer into the drift cell and hence weakened signal, and higher IE will cause ions to traverse too deep into the

cell before they start to drift, thus lowering the effective separation length. The ions drift within the cell in the uniform linear electric field which is maintained between the cell entrance and exit. The ions exit the cell in order of their mobility *K*, and are re-focussed in the post-cell hexapole and proceed to the ToF region to be mass-measured. The next 40-μs injection pulse will take place after 200·ToF period milliseconds, thus giving the value of pre-injection accumulation time in the hexapole. The ToF period, in its turn, is directly proportional to the *m*/*z* acquisition range. For example, to analyse ions up to 3400 *m*/*z* will result in maximum flight time of 100 μs. Thus, any two mobility events are separated by 200·100 μs = 20 ms or, in other words, take place at a frequency of 50 Hz ($\nu = 1/t = 1/20$ ms = 50 Hz). This frequency is set by the user on the digital delay/pulse generator DG535 (Stanford Research Systems, Sunnyvale, CA, USA).

The output of the mobility experiment is the TIC featuring a peak (or closely 'eluting' peaks, with resolution of 20) termed 'total arrival time distribution' (tATD). Typically 11 such peaks are acquired to be summed to yield greater signal intensity and to compensate for the poor sampling rate due to the innate periodicity of the ToF analyser. The MassLynx 4.1™ (Waters, Manchester, UK) built-in software was modified to accommodate for the ion mobility capacity, and allows the user to select a specific *m*/*z* (or a range of *m*/*z*) and generate a selected ion ATD.

2.2.2 Typical workflow of IM-MS data acquisition and analysis

A typical data acquisition procedure involves performing a number of mobility events at different drift voltages applied across the cell. The voltages range from 60 to 10 V in steps of 10 or 15 V, e.g. 60, 45, 35, 30, 25, 20, 15 and 10 V, and extractor voltage is adjusted by the same number of volts as the drift cell potential. Resulting ATDs are recorded along with pressure and temperature for each of the voltages. The use of multiple voltages instead of one removes the necessity to measure the instrumental dead time t_0 for each ion, as will be explained below.

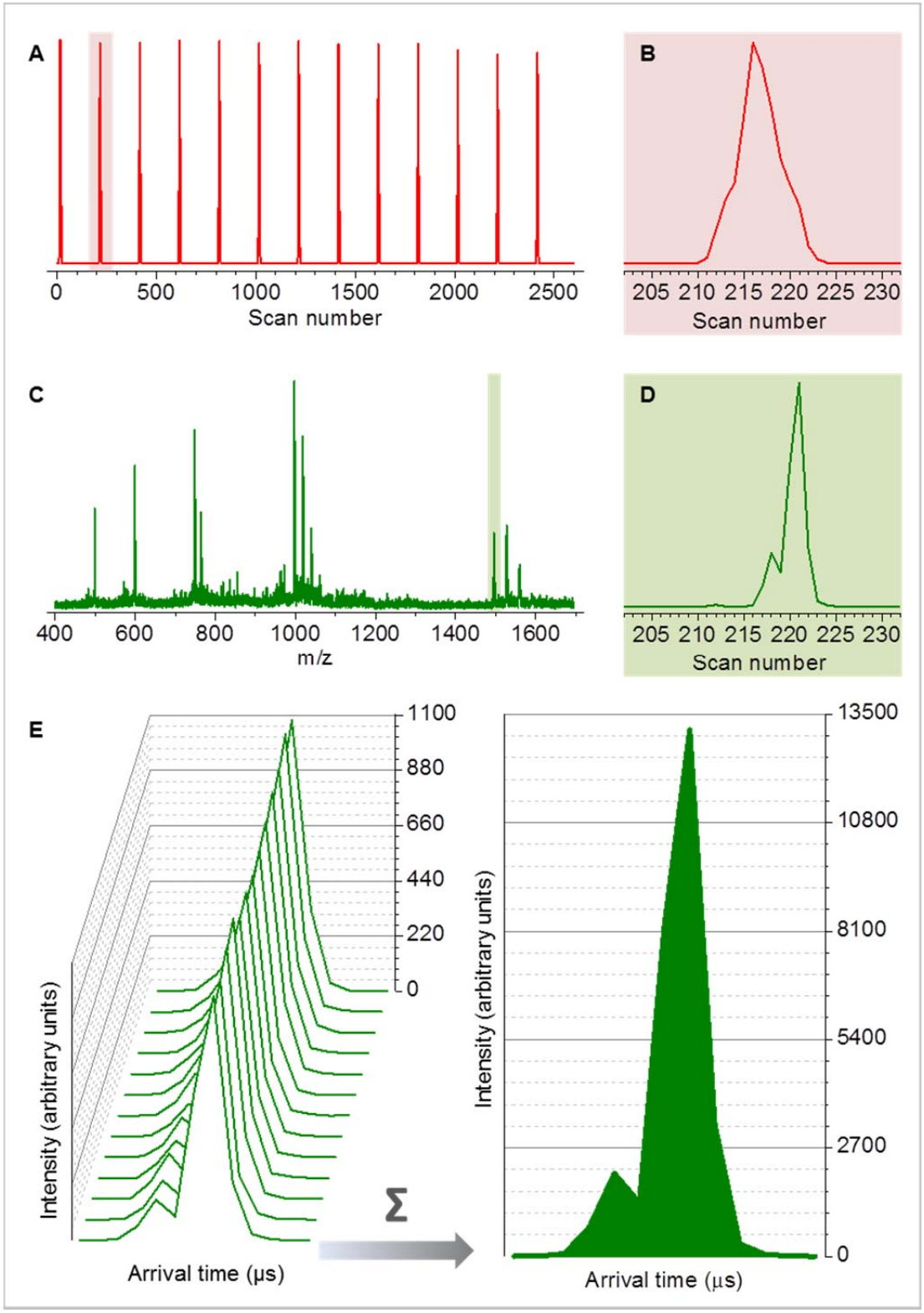


Figure 2-3 *Output of a mobility experiment and first data processing steps. (**A**) – TIC of raw ATDs at drift potential 60 V; (**B**) – a zoom of the peak highlighted in (A); (**C**) – full mass spectrum of ions arriving at the detector after mobility separation; (**D**) – XIC of the 1496 m/z peak highlighted in (C); (**E**) – illustration of summation process of XIC peaks at 1496 m/z.*

Figure 2-3A shows the output of the raw data collected for 13 identical mobility experiments carried out at one of the drift voltages (60 V). The pale pink box highlights the one of the ATDs, and the zoomed in peak is shown in ***Figure 2-3B***. By combining all the scans under that peak generates full mass spectrum (***Figure 2-3C***) of all ions that reached the detector. For any *m*/*z* value (or a range of *m*/*z* values), an extracted ion chromatogram (XIC) can be reconstructed by combining the *m*/*z* range in question, as shown by the pale green box across the ion at 1496 *m*/*z*. The resulting XIC is shown in ***Figure 2-3D*** and features two closely 'eluting' peaks. The ATDs on panes ***B*** and ***D*** are aligned by their scan number to highlight the relative position of the XIC peak with respect to the TIC. In this example, intensity of the XIC is 3% of that of the TIC. As has been mentioned in section ***2.2.1***, the data quality is enhanced by summing the repeated ATDs, using a simple Excel spread-sheet template. ***Figure 2-3E*** shows the 13 ATDs, all at approximate intensity of 1,100 counts, which after summation yield a single ATD peak (***Figure 2-3F***) with intensity increased by almost a factor of 13 (13,500 counts).

The next step in ion mobility data analysis is to obtain arrival times of the ions of interest, which involves combining the results for each of the drift voltages. ***Figure 2-4A*** shows ATDs obtained at several drift potentials arranged in order of decreasing drift voltages. The scan numbers are converted into arrival times by multiplying the former by the ToF pusher period to yield values in microseconds. As has been shown above (***Figure 2-3D***), an ATD at a single voltage can feature one or more peaks, whose relative ratio may either remain constant or change with changing voltage. In this example, the latter trend is observed, which indicates that the earlier-arriving ions are the dimer of the later-arriving species. If the relative peak ratio remained constant across all voltages, the two resolved species would

be conformers of the same ion. The increased proportion of the dimer (or other multimer) at decreasing drift voltages arriving at the detector is due to fact that these species have higher charge. Higher charged species are better focussed by the confining voltages and less susceptible to radial diffusion during their drift, the effect that is more pronounced at lower drift voltages as the time spent in the cell is increased.

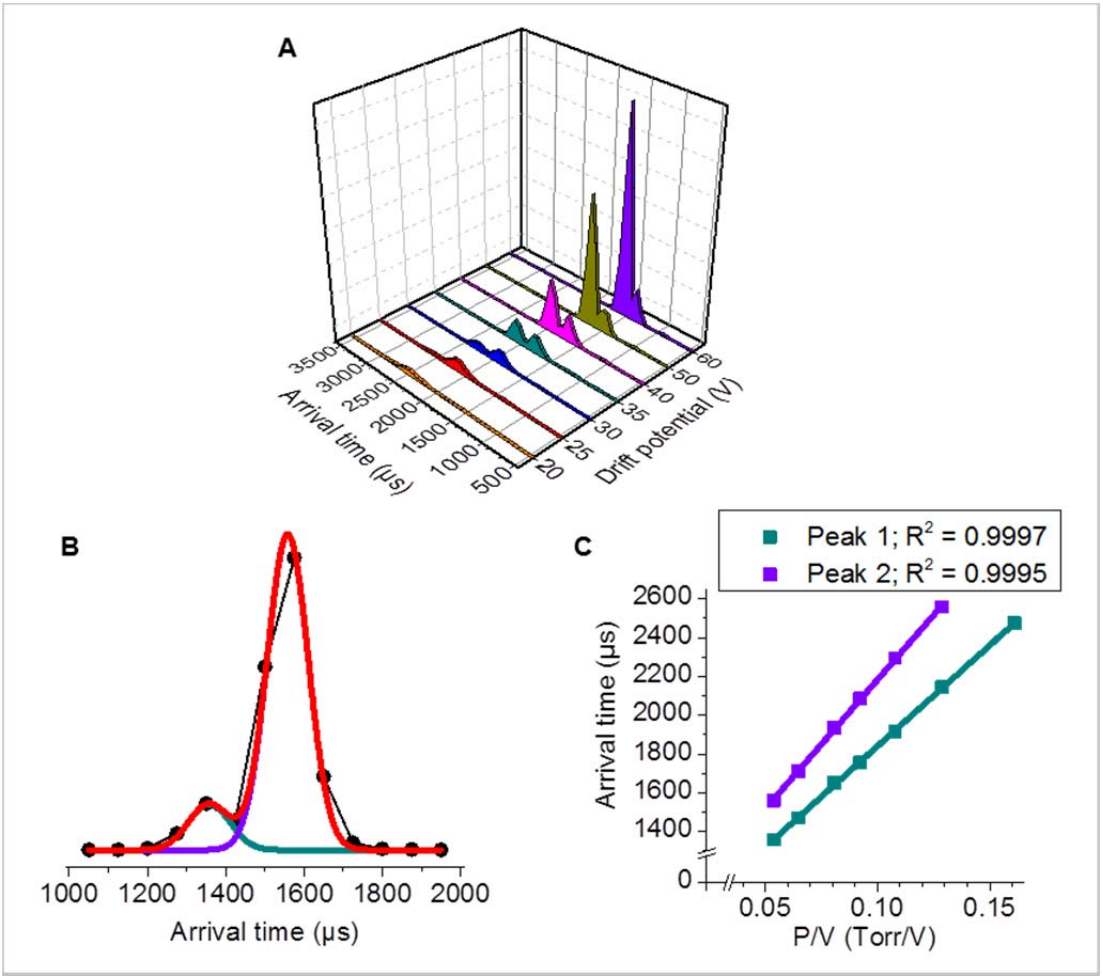


Figure 2-4 *Further steps of data processing from ATDs to mobility plots. (**A**) – a set of XICs for the selected m/z value obtained at discrete drift potentials; (**B**) – Gaussian fitting to the ATD data points to elucidate average arrival times for populations of ions; (**C**) – mobility plots created from arrival times at discrete drift voltages, from which mobility values are calculated.*

For each peak on these ATDs, a Gaussian distribution is fitted to the experimental data bins, along with their cumulative curve, using Origin 8.5.1 graphing software (OriginLab Corporation, Northampton, MA, USA), as shown in ***Figure 2-4B***. Maxima of the resulting peaks are used as arrival times t_a for each of the resolved species, and their values are plotted as a function of their corresponding experimental P/V ratios. As has been shown in section ***1.3.2***, mobility K is often expressed through reduced mobility K_0 to decouple it from the instrumental platform and experimental conditions:

$$K = K_0 \frac{P_0 T}{T_0 P}, \qquad \textbf{Equation 2-1}$$

where T_0 = 273.15 K is the standard temperature and P_0 = 760 Torr is the standard pressure.

On the other hand, mobility K can be expressed as:

$$K = \frac{L^2}{t_d V} \qquad \textbf{Equation 2-2}$$

By combining the two equations, the drift time can be expressed:

$$t_d = \frac{L^2 T_0}{K_0 T P_0} \frac{P}{V} \qquad \textbf{Equation 2-3}$$

This equation describes a straight line with the slope M inversely proportional to the reduced mobility:

$$M = \frac{L^2 T_0}{K_0 T P_0} = \frac{L^2}{KP} \qquad \textbf{Equation 2-4}$$

Thus, a straight line can be fitted to the experimental arrival time data points (***Figure 2-4C***) with its y-intercept equal to the instrumental dead time t_0. The line fitting quality is assessed by the variance parameter R^2, and the data whose fits do not reach the criteria of 0.999 are discarded.

Thus, mobility K can be calculated without the need to determine the dead time t_0, as the parameters in ***Equation 2-4*** are either a known fixed parameter (L) or measurable experimental parameters (K and P). From the values of reduced mobility K_0, rotationally averaged collision cross sections Ω (reported in $Å^2$) of analyte ions are calculated:

$$\Omega = \frac{3ze}{16N}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{1}{K} \qquad \textit{Equation 2-5}$$

Where z is the charge number, e [C] is the elementary charge, N [m^{-3}] is the helium number density, μ [kg] is the reduced mass of the ion-neutral pair, T [K] is the drift gas temperature, K [$m^2 \cdot V^{-1} \cdot s^{-1}$] is the mobility.

The drifts in pressure during the mobility experiments are in the range of 0.01 Torr, whereas variations in temperature are more profound when conducting the low- and high-temperature work and can reach 0.5 K. ***Table 2-2*** summarises average deviations in pressure and temperature as a function of cell temperature. To minimise this effect, a finer controlled cell heating system and an automated feedback loop for the cooling system is recommended, along with better insulation of the liquid nitrogen Dewar to minimise thermal loss.

Cell temperature setting, K	Average deviation:	
	of temperature, K	of pressure, Torr
360	0.5	0.010
300	0.2	0.011
260	0.4	0.009

Table 2-2 *Average deviations in pressure and temperature as a function of cell temperature setting. The values are calculated as the average of the absolute deviations of data points from their mean.*

2.2.3 Typical MoQ-ToF settings

Typical instrument settings used in the presented work are shown in ***Table 2-3***.

Parameter	Value	Comment
Capillary (kV)	1.50	
Cone	50 – 70	
Extractor*	100	137 for the study in ***Chapter 5***
RF Lens	0.75	
Source Temp (°C)	80	
LM Resolution	5	
HM Resolution	5	
Collision Energy	3.8	
Ion Energy	1.7	
Steering	–0.36	
Entrance	37	
Pre-filter	7.6	
Transport	3.6	
Aperture2	12.2	
Acceleration	200	
Focus	5	
Tube Lens	85	
Guard	47.7	
TOF (kV)	7.20	
Reflectron	35.00	
Pusher Cycle Time (μs)	70 – 110	Adjusted according to *m*/*z* range
Multiplier	654	
MCP	2400	3000 for the study in ***Chapter 5***
Injection Energy	33-38	Up to 58 for the study in ***Chapter 5***
P(He) in drift cell (Torr)	3.2-3.5	
T(He) in drift cell (K)	300	260, 300 & 360 for the study in ***Chapter 5***

Table 2-3 *Typical MoQ-ToF settings. Values are shown in volts, unless stated otherwise. (*) Extractor voltage is given at cell drift potential 60 V; this value is decreased by the same number of volts as the drift cell during a mobility experiment.*

2.3 Peptide and protein samples

2.3.1 Synthetic peptides

All of the peptides reported in this work (vCP1, ZiCop and Pp) were synthesised using standard solid-phase Fmoc-based technique using rink-amide resin and HBTU activation. They were supplied as aqueous solutions by the Woolfson group (University of Bristol). Peptide vCP1 was synthesised using a Liberty CEM microwave synthesiser and was N-terminally acetylated and C-terminally amidated. Peptides ZiCop and Pp synthesis was carried out on a PS3 peptide synthesiser (Protein Technologies Inc., Tucson, AZ, USA). All peptides were purified by HPLC (JASCO, Great Dunmow, UK) on C18 columns, and their identity confirmed by MALDI-ToF mass spectrometry using α-cyano-4-hydroxycinnamic acid (CHCA) matrix (ToF/ToF™ 4700 Proteomics Analyser, Applied Biosystems, Paisley, UK). Protein concentrations were measured on a PerkinElmer Lambda-25 spectrophotometer, and determined using Beer's law (University of Bristol) and confirmed at the University of Edinburgh using micro-volume UV-visible NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). On arrival, peptides were stored at 4°C in their aqueous solutions, and prior to all MS and IM-MS studies all cysteine-containing peptides were incubated for ~1 min with a 10-fold molar excess of TCEP to selectively reduce any potential disulphide bond formation. Peptides were made up to concentrations of either 20 or 50 μM in 20 or 10 mM ammonium acetate buffer with pH adjusted to either 6.8 or 7.2 with additions of either ammonia

or acetic acid. All solutions were made up to contain 5% IPA to aid electrospray desolvation.

2.3.2 Lysozyme

Lysozyme from chicken egg white (product number L6876, ≥90% purity, 14,307.16 g/mol) was purchased from Sigma-Aldrich Company Ltd (Dorset, UK). 1 mM stock solution of the protein was prepared in deionised water and the aliquots were stored at –20°C. For CCS measurements, 50 μM concentrations of lysozyme were prepared in deionised water with addition of 40-fold molar excess of sodium iodide (2 mM).

2.3.3 Cytochrome c

Cytochrome *c* from equine heart (product number C7752, ≥95% purity, 12,359.94 g/mol) was purchased from Sigma-Aldrich Company Ltd (Dorset, UK). 1 mM stock solution of the protein was prepared in deionised water and the aliquots were stored at –20°C. For CCS measurements, 50 μM concentrations of cytochrome *c* were prepared in deionised water with addition of 40-fold molar excess of sodium iodide (2 mM).

2.3.4 Bovine pancreatic trypsin inhibitor

Bovine pancreatic trypsin inhibitor, BPTI (product number T0256, essentially salt-free, 6,517.49 g/mol) was purchased from Sigma-Aldrich Company Ltd

(Dorset, UK). 1 mM stock solution of the protein was prepared in deionised water and the aliquots were stored at –20°C. For CCS measurements, 50 μM concentrations of BPTI were prepared in deionised water with addition of 40-fold molar excess of sodium iodide (2 mM).

2.4 Metal salts and reducing agent for metal-binding studies on synthetic peptides, and salt for protein-anion interaction study

Peptide-metal binding studies were carried out on four metals, which were used from their acetate salts without further purification. 1 mM aqueous stock solutions were prepared and stored at 4°C. Cobalt, calcium and copper acetates (CoAc: anhydrous, 177.02 g/mol, 99.995% purity, product number 399973; CaAc: monohydrate, 176.19 g/mol, ≥99% purity, product number C8570; CuAc: anhydrous, 181.63 g/mol, 99.999% purity, product number 517453,) were obtained from Sigma Aldrich (Dorset, UK). Zinc acetate (ZnAc: dihydrate, 219.51 g/mol, product number Z20-500,) was obtained from Fisons (Loughborough, UK). Sodium iodide (product number S/5120/50, ≥99% purity, 149.89 g/mol) was obtained from Fisher Scientific (Loughborough, UK). Stock solution of NaI was prepared in deionised water and used without further purification. TCEP hydrochloride (286.65 g/mol, product number 20490) was acquired from Pierce Biotechnology (Loughborough, UK) and 1 mM aqueous stock was prepared and stored at 4°C.

2.5 Solvents and ammonium acetate buffer

LC-MS grade methanol (32.04 g/mol, product number A456-212) and HPLC grade isopropyl alcohol (60.10 g/mol, product number P/7508/17) were obtained from Fisher Scientific (Loughborough, UK). Water was deionised on Sartorius arium® 611, a four-stage purification process to yield purity in the resistivity range of 18.0 – 18.2 MΩ·cm. This system uses mixed bed resins and activated carbon in conjunction with an ultrafilter and a 0.2-micron final filter to further purify water pre-treated by distillation, deionization and reverse osmosis. Ammonium acetate (77.08 g/mol, ≥99% purity, product number A/3440/50) was purchased from Fisher Scientific (Loughborough, UK) and fresh aqueous solution stocks were prepared weekly at pH 6.9. Sample pH in ammonium acetate buffer was adjusted by additions of ammonia solution (17.03 g/mol, 0.88 S.G. 35% NH_3, product number A/3280/PB17, Fisher Scientific, Loughborough, UK) to raise pH and acetic acid (60.05 g/mol, 99.9% purity, product number 20104.334, VWR, Leicestershire, UK) to lower pH of the sample.

2.6 Data acquisition and analysis

2.6.1 Calculation of dissociation constants

Calculation of K_d using gas-phase titrations of the peptide with one ligand, along with two-ligand competitive binding experiments, uses the approach described by Sannes-Lowery and co-workers[1] and theory described by van

Holde *et al.*[2] Using nESI-MS, the stoichiometry and relative abundance of the apo-peptide and the peptide-metal complex can be determined simultaneously *via* their first monoisotopic peak intensities.

When a protein has *n* ligand-binding sites, the reaction is then:

$$P + nL \Leftrightarrow PL_n$$ **Equation 2-6**

When a protein has one ligand-binding site, the reaction above will take form:

$$P + L \Leftrightarrow PL$$ **Equation 2-7**

In this case, the dissociation constant K_d is expressed as follows:

$$K_d = \frac{[P][L]}{[PL]}$$ **Equation 2-8**

where [P] is the concentration of the unbound state of the peptide, [L] is the concentration of free ligand and [PL] is the concentration of bound ligand.

Initial concentrations of peptide [P_i] and added metal salt – ligand [L_i] are always known. The degree (*i.e.* velocity v) of binding can be defined as the average number of bound ligand units per target unit and is expressed as:

$$v = \frac{[PL]}{[P] + [PL]}$$ **Equation 2-9**

Therefore the sum in the denominator of the ***Equation 2-9*** is the total concentration of the peptide added initially:

$$[P_i] = [P] + [PL]$$ ***Equation 2-10***

The concentration of the free ligand is given by the difference between initial (added) and bound ligand concentration:

$$[L] = [L_i] - [PL]$$ ***Equation 2-11***

When the normalised values obtained from ***Equation 2-9*** are plotted against those in ***Equation 2-11***, and the Hill curve is fitted to the data points, the point on this curve at 0.5 binding (or 50% target saturation) gives the dissociation constant in ligand concentration units. The Hill curve is described by the following sigmoidal equation:

$$v = \frac{V_{\max}[L]^n}{K_d^n + [L]^n}$$ ***Equation 2-12***

where V_{max} is the highest degree of binding (maximum velocity achieved by the system), n is the number of cooperative sites and K_d is the Michaelis constant that is equal to the ligand concentration at which the binding rate is half of V_{max}. In case of single-site binding (n = 1), no cooperativity is observed, and the curve becomes hyperbolic and is described by the Michaelis-Menten equation:

$$v = \frac{V_{\max}[L]}{K_d + [L]}$$ ***Equation 2-13***

The linear part of the slope of either of such curves gives a clear indication of the range of concentrations at which dissociation occurs.

2.6.2 Protein-anion interaction study

2.6.2.1 Ion source temperature effect

To decouple the protein-salt clustering events from the gas-phase 'artefacts', the source temperature effect study has been conducted. For each cell temperature setting, three different source temperatures were used (60, 80 and 100 °C) to obtain mass spectra of the proteins with salt, and the relative ratios of the free and salt-bound ions were compared. This was carried out in 3 experimental repeats taken on different days, and the average of 3 mass spectra for each source temperature was reported.

2.6.2.2 Ion arrival times measurements

Collision cross sections of the ion series of BPTI, cytochrome *c* and lysozyme were measured in triplicates and the mean values reported. The data were collected at three different drift cell temperatures for each of the proteins – 260 K ('cold'), 300 K ('ambient') and 360 K ('hot'); the ion source temperature setting was 80 °C. Resulting arrival time distributions presented themselves as complex profiles often featuring more than one coeluting peak (usually two, with tailing and fronting), making the assignment of the peak maxima ambiguous and CCS calculation unreliable. Also, assigning one or two discrete CCS values to a vastly populated conformational domain would not reflect the physical reality of the events in the drift cell, as the results would be 'over-interpreted'. Therefore, the data need to be presented in their more 'native' form, *i.e.* as ATD profiles (similar to the one shown in ***Figure 2-4B***), and at the same time, a comparison of the results obtained at different experimental conditions has to be possible. To enable this, ion arrival time t_a

scale needs to be converted to CCS scale (*via* drift time t_d), for which the dead time value t_0 has to be known. The strategy described in section **2.2.2** cannot yield the latter, so it had to be modified. The new method of Gaussian curve fitting has been developed with the kind help from Dr Jason Kalapothakis to afford simultaneous fitting of the whole drift voltage series for a particular ion at once (previously, each drift voltage was fitted separately). Apart from the improved data workup efficiency, this method also: (i) takes into account the inter-dependency of the ATD parameters within one drift voltage series; (ii) yields the dead time value t_0 associated with each of the m/z values, enabling the direct use of the ***Equation 2-2*** to obtain mobilities, and hence CCS, at known drift voltages. The Origin graphing software enables the user to define their custom functions, therefore the existing Gaussian fitting function has been modified to enable fitting the two-ion flux at specified drift voltage series. Resulting values of dead time t_0 and were used to convert arrival times into CCS using ***Equation 2-2*** and ***Equation 2-5***. All the data obtained at 300 K were treated using this method, and the resulting mean values of t_0 were used to process the rest of the data obtained at 260 and 360 K without curve-fitting, as the t_0 for any given m/z value is independent of the drift cell temperature.

2.7 Molecular modelling

Force field measurements for the peptide systems (vCP1 and ZiCop) were carried out using the Amber10 modelling package implementing Amber99 force field[3] which contains the divalent zinc atom parameters. For structure minimisation of the proteins (lysozyme, cytochrome *c* and BPTI) the module

sander of Amber11[4] with Amber99SB-ILDN force field[5] was used. The ionisation state of all the peptides and proteins was assigned as it would be in solution at pH 7.00 using the H++ algorithm which automates the procedure of protonating the side chains of the residues, according to the specified value of the solvent pH [6;7].

2.7.1 vCP1 system

For the holo-vCP1, a homology model of the peptide based on the consensus zinc finger structure was submitted to the H++ algorithm. The pH value was set to 7.00, to obtain a system total charge of +3.0. The model adopted to simulate the Zn-ligand interaction was non-bonded[8;9], which entertains only electrostatic and van der Waals' forces between the protein and prosthetic groups. After minimising the peptide and subsequently equilibrating it at 300 K, a production run of 10 ns was carried out in canonical ensemble, with a time step of 0.5 fs.

For the apo-vCP1 structure, a protonation state found for pH 7.00 was used, which resulted in a total charge of +3.0 (as observed experimentally), with sulphur atoms on the cysteines hydrogenated, in order to represent the fully reduced state of the system. Then two strategies were followed:

❖ A simulated annealing (SA) protocol starting from a fully extended peptide was applied. The protocol repeats an energy minimisation n (= 100) times, followed by "hot" dynamics at 800 K and subsequent stepwise cooling of the system until 0 K. For every minimised structure the TM collision cross section was calculated (*via* MobCal). The average TM collision cross for these

100 structures Boltzmann weighted by their energy was obtained. To obtain a time evolution of the collision cross section, MD simulations were performed on two representative structures: the first is the one with the lowest energy value and the second is the one whose TM collision cross section value is the nearest to that experimentally measured. These two structures were heated up to room temperature and then NVT dynamics up to 5 ns was followed, with a time-step of 0.5 fs. Their time averaged TM collision cross sections were then calculated with MobCal.

❖ In the second strategy the zinc atom was removed from the holo structure and an MD simulation run for to 10 ns. Whilst strategy 1 is probably a better method by which to sample the gas phase conformations that are observed experimentally, to be consistent with the method taken for the holo-peptide, ***Figure 3-5*** in ***Chapter 3*** contains a representative structure from the second strategy. The CCS measured for conformations in each case are indistinguishable.

2.7.2 ZiCop system

The time-step for the simulations *in vacuo* was 0.5 fs (as no constraints were applied). Following minimisation and thermalisation of the structures at 300 K, the production runs were carried out in canonical ensemble for up to 10 ns. The non-bonded[8] model was adopted to simulate the interaction of the Zn ligand with ZiCop, which uses only electrostatic and van der Waals forces. The resulting structures of ZiCop and Zn-ZiCop were subjected to simulated annealing to probe a wider conformational space for the candidate geometries. The lowest energy structure found within the band of

experimental cross section values was selected to calculate MD collision cross sections in MobCal program[10; 11] using trajectory method.

2.7.3 Lysozyme, cytochrome c and BPTI

Experimental CCS of the protein ions were compared to the published[12] X-ray crystallographic data by plotting the values of the cross sections of the latter calculated from the coordinates in PDB[12] files. The following PDB files were used: 3AW6 (for lysozyme), 1HRC (for cytochrome *c*) and 1BPI (for BPTI), and the charge states were assigned as 8+, 6+ and 6+, respectively by the H++ server imposing pH 7.00. Each protein structure was minimised in the gas phase at three temperatures – 260 K, 300 K and 360 K using an "infinite" radial cut-off (r_c = 999 Å) for the non-bonded interactions. CCS were calculated using trajectory method (TM) as implemented in MobCal[10; 11] program and at the same temperatures at which the measurements were taken (260 K, 300 K and 360 K). The error associated with such calculations is below 0.5% for lysozyme and cytochrome *c*, and ~1% for BPTI (estimated from previous repeat calculations, not shown).

2.8 References

1. Sannes-Lowery, K. A., Griffey, R. H. & Hofstadler, S. A. (2000). Measuring Dissociation Constants of RNA and Aminoglycoside Antibiotics by Electrospray Ionization Mass Spectrometry. *Anal. Biochem.* **280**, 264-271.

2. van Holde, K. E., Johnson, W. C. & Ho, P. S. (1998). *Principles of Physical Biochemistry*, Prentice-Hall, Inc.

3. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B. & Woods, R. J. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668-1688.

4. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A. & Kollman, P. A. (2010). *AMBER 11*, University of California, San Francisco.

5. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* **78**, 1950-1958.

6. Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S. & Onufriev, A. (2005). H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **33**, W368-W371.

7. Myers, J., Grothaus, G., Narayanan, S. & Onufriev, A. (2006). A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins Struct. Funct. Bioinf.* **63**, 928-938.

8. Sakharov, D. V. & Lim, C. (2005). Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **127**, 4921-4929.

9. Bredenberg, J. & Nilsson, L. (2001). Modeling zinc sulfhydryl bonds in zinc fingers. *Int. J. Quantum Chem.* **83**, 230-244.

10. Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C. & Jarrold, M. F. (1996). Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. *J. Phys. Chem.* **100**, 16082-16086.

11. Shvartsburg, A. A. & Jarrold, M. F. (1996). An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86-91.

12. http://www.rcsb.org/pdb/home/home.do. RCSB Protein Data Bank.

Chapter 3

Ion mobility mass spectrometry as a tool for protein design. Part 1: a case study on consensus zinc finger peptide

3.1 Introduction

This chapter describes characterisation of a consensus zinc finger peptide using ion mobility mass spectrometry and molecular dynamics simulations. The system considered here is a synthetic consensus zinc finger sequence (vCP1) that is responsive to zinc: it adopts a zinc finger fold in the presence of Zn^{2+} by coordinating the metal ion by two cysteines and two histidines. This peptide has been selected as a reference for the zinc-bound state and a simple model to refine the characterisation method in preparation for analysis of a more sophisticated dual switching system that is capable of adopting either of the two conformations in response to a stimulus:

zinc finger or coiled coil. This latter system will be discussed in the following chapter.

Although new evidence suggests that *the* native protein fold really exists as an ensemble of conformations in dynamic equilibrium[1], the fundamental tenet that primary amino acid sequence of a given protein defines its subsequent folded structure still holds true. The scientific community has been intrigued by the prospect of being able to predict protein function from its sequence[2; 3]. Understanding of the sequence-to-structure relationships in proteins can give an insight into a fundamental question of how proteins fold from an amino acid chain to a biologically active protein. A systematic approach has been taken by a number of research groups[4; 5; 6; 7; 8; 9] to get an insight into how the protein fold is defined by its primary amino acid sequence.

An important practical aspect of understanding sequence-to-structure relationships includes, but certainly is not limited to, production of novel bio-orthogonal systems that can find their application in new efficient therapies, affinity-based targeted drug delivery, bio-sensors, *etc*[10; 11; 12]. There are two approaches that are generally adopted to elucidate links between protein sequence and structure. The first approach is called the 'protein folding problem', and it tries to predict a protein's fold from its amino acid sequence (engineered proteins). The second strategy is referred to as the 'inverse protein folding problem', also known as 'protein design', and it attempts to identify amino acid sequences that adopt a certain fold[13] (synthetic proteins).

Synthetic and engineered proteins also present appealing model systems for studying more complex biological events through mimicry. This task requires an in-depth understanding of underlying principles of how protein function can be related back to its primary sequence *via* its conformation[14]. Both synthetic and engineering approaches use chimera proteins, however they differ in the source of the material[15]. Engineered proteins owe their existence to the refinement of gene manipulation techniques that provide mutated proteins with the desired properties and functions. Synthetic proteins, on the other hand, are products of *de novo* design, which in itself is a more ambitious venture aiming at creating completely new protein structures with specific pre-defined and novel properties[16]. Despite the complexity of the task at hand, a number of attempts at protein design have proven successful[17; 18; 19; 20].

Traditionally, protein design relies heavily on solution phase methods for characterisation, for example circular dichroism and UV-visible spectroscopies. The benefits and drawbacks of these methods were outlined in section ***1.4.1***. Here, mass spectrometry and ion mobility mass spectrometry coupled with molecular dynamics (MD) simulations are used to probe interactions between divalent metal ions and a synthetic Cys_2His_2 zinc finger peptide as an alternative characterisation route. This work expands the analytical toolbox available for protein design, thus enhancing characterisation capabilities.

The system under investigation is a consensus zinc finger peptide vCP1 (variant consensus peptide 1) that forms the first domain of the three-zinc-finger protein (theoretical $MW_{average}$ = 10,104.2157 Da) designed and characterised by the Berg group[21]. The structure of this protein interacting

with an oligonucleotide corresponding to a favourable DNA binding site has been solved by X-ray crystallography. In this work, gas-phase methods are used to characterise the following aspects of the vCP1 system:

- binding specificity of Zn^{2+} and Co^{2+};
- binding selectivity of these two ions compared to Ca^{2+} and Cu^{2+};
- conformational change as a result of Zn^{2+} and Co^{2+} binding and structural stability arising from this;
- molecular dynamics simulations were used to gain an atomistic level insight into the vCP1 behaviour.

3.2 Design of the vCP1 system

The synthetic peptide for a consensus zinc finger sequence, vCP1, has been derived from the work of Berg *et al.*[21] Their design featured the most prevalent amino acid at each position of the sequence alignment. This sequence was found to fold in the same manner as natural zinc fingers, and bind zinc more tightly than any natural peptide tested to date at the time[22]. The same group investigated the effects of creating a minimalist zinc finger-like peptide in which all non-conserved positions were alanine (a small "inert" amino acid) and lysine (to aid solubility). Remarkably, the NMR methods used to probe the three-dimensional structure of this peptide revealed that the peptide folded very much like a natural zinc finger[23]. This

suggested that the domain was robust and simple enough in nature to be amenable to design using principles from structure prediction. These were primary reasons to select vCP1 as a model system to refine gas-phase analytical method. The peptide vCP1 comprises just 26 amino acids making it a good model for studying peptide-metal interactions in solution and the gas phase. ***Figure 3-1A*** shows the sequence of vCP1 (top string), with the zinc-binding residues are highlighted in green (Cys) and turquoise (His), and the three other highly conserved residues – tyrosine, phenylalanine and leucine, are shown in purple[24]. vCP1 was obtained by sequence search in the Protein Data Bank (PDB) against the sequence shown at the bottom string of ***Figure 3-1A***. The well-established coordination of zinc ion by Cys_2His_2 tetrahedral geometry[24] is shown in the equilibrium schematic below. ***Figure 3-1B*** summarises some basic facts about the vCP1 sequence. To probe the affinity of this peptide for metal ions the Berg group performed several different spectroscopic absorption assays[22], two of which are highlighted here, as they are most relevant to this work. The first one involved direct titration of the peptide with Co^{2+} and it produced a curve that was fit with a dissociation constant K_d < 0.1 μM. The second method used Zn^{2+}/Co^{2+} competition and the K_d was determined to be ~2 pM. One-dimensional NMR spectra of the peptide in the presence on Zn^{2+} at pH 7.54 revealed that only fully folded form was present. These finding suggest that in the presence of zinc ion the binding equilibrium is shifted to the right.

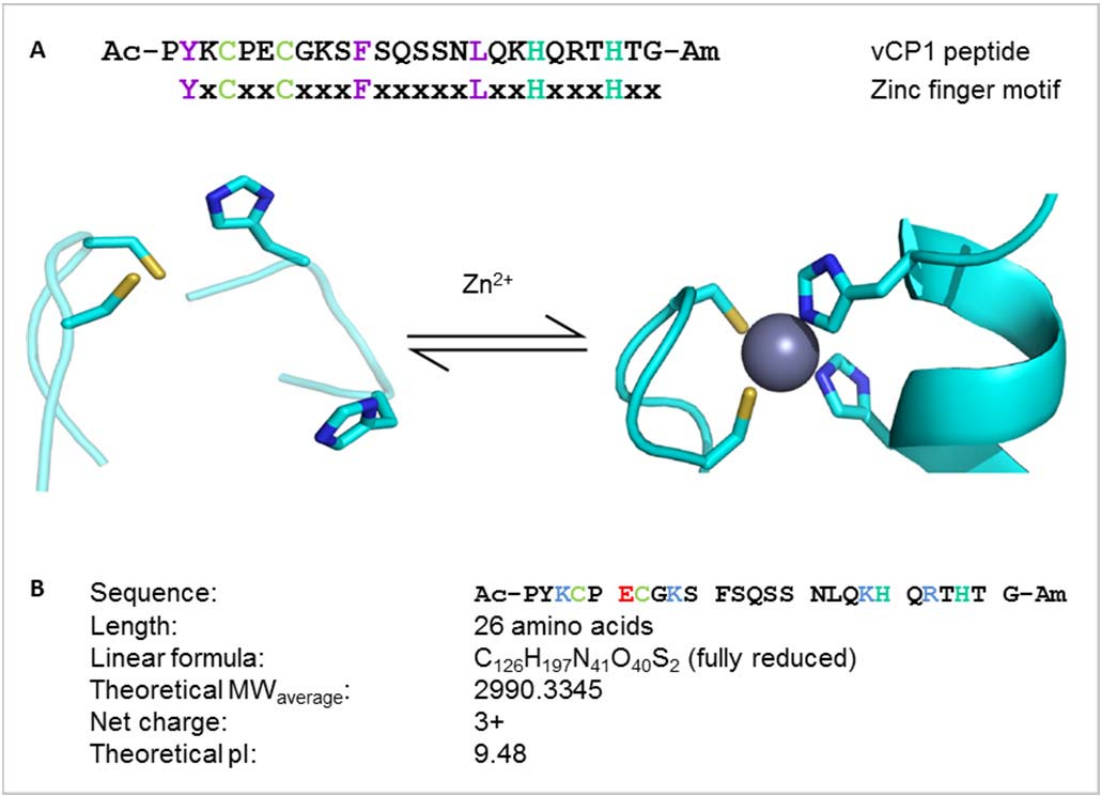


Figure 3-1 *The vCP1 peptide. (**A**) Sequence, design criteria and equilibrium in the presence of zinc ion. The sequences show key cysteine and histidine residues in green and turquoise, respectively. Other conserved residues form a small hydrophobic core and are shown in purple. The schematic shows the proposed rearrangement of zinc-binding residues of the zinc finger peptide when the two cysteine (sulfhydryl groups in yellow) and two histidine residues (imidazole nitrogen atoms in blue) coordinate Zn^{2+} (slate grey)[25]. Adapted from PDB entries 1Z60 and 2DRP by Dr Craig Armstrong, University of Bristol. (**B**) Summary of facts relating to vCP1. Acidic and basic amino acid residues in protein sequences are shown in red and blue respectively. Cysteine and histidine residues in green and turquoise, respectively.*

The work presented in this chapter proposes a gas-phase platform to characterise *de novo* designed proteins in general and metal binding in particular as both a competitive and complementary strategy. To probe selectivity and specificity of vCP1 for different metals, apart from Zn^{2+}, a range of other divalent metal cations were used: Co^{2+}, Cu^{2+} and Ca^{2+}. The first three are transition metals and the last one is alkaline earth metal, whose selected physical properties are summarised in ***Table A 2*** of ***Appendix***.

3.3 Overview of interactions – stoichiometry of binding

The equilibrium described in ***Figure 3-1A*** was probed by nESI-MS following the experimental procedure described in ***Chapter 2***, ***Table 2-1***. The resulting mass spectra (***Figure 3-2A***) revealed that the vCP1 peptide presents two major charge species assigned as $[vCP1+3H]^{3+}$ and $[vCP1+4H]^{4+}$, with the former being the most dominant, the latter is approximately half its intensity. A very low-intensity $[vCP1+2H]^{2+}$ signal is also observed. Mass spectra of vCP1 with additions of equimolar amounts of metal acetate salts (***Figure 3-2B*** through ***E***) show that extent of binding is varied for different metals.

Surprisingly, Zn does not bind to vCP1 in a 1:1 ratio, leaving approximately one half of its population in the apo-form (***Figure 3-2B***). The zinc-bound 4+ charge state constitutes only a fraction of its apo form signal. For all spectra, impurities originating from the metal salt are denoted by black asterisks, and TCEP (tris(2-carboxyethyl)-phosphine) adducts[26] are shown in orange asterisks. Closer analysis of spectral ranges with data around the orange asterisks revealed (not shown) that TCEP associates predominantly with the metal-bound vCP1. This association presumably takes place between the carboxylic groups of TCEP and the positively charged amino acid residues of vCP1 – Lys and Arg as they all face outward in the metal-bound conformation (as is revealed by MD simulations discussed below in section ***3.5***). There is a higher probability of TCEP being bound to Arg than Lys, as the pK_a of Arg is 13.2 (compared to the lower value of 10.3 for Lys). The analysis of calcium interacting with vCP1 is a notable exception, as the strongest signal of TCEP association with vCP1 is the one without the metal

(perhaps not surprisingly, as the intensity of signal from the holo-form here is very low).

Cobalt binds to vCP1 at almost 1:1 stoichiometry, indicating higher selectivity for this metal compared to zinc (***Figure 3-2C***). Interestingly, when the peptide binds both zinc and cobalt it only presents a single charge state following ESI-MS, suggesting that in solution the fold is stable, with a single configuration of solvent-exposed residues giving rise to one charge state.

Figure 3-2D shows the mass spectrum of vCP1 and calcium acetate mixed in equimolar amounts, which is strikingly similar to the mass spectrum of the apo-vCP1 (*cf*. pane ***A***). Very little binding occurs, indicating the non-specific nature of it. Indeed, as a divalent cation from group II of the Periodic Table, calcium is a 'hard' metal ***(Table A 2*** in ***Appendix***) that binds to carboxyl groups of Asp and Glu[27]. vCP1 provides only one such site – the Glu-6 residue.

Figure 3-2E features the vCP1 interaction with copper ion, where only a small fraction of the peptide is metal-associated. A feature unique to this mass spectrum, compared to all of those discussed above, is that the metal-free 2+ charge state gains prominence here. Both calcium- and copper-bound peptide do not feature the holo-4+ ion, in contrast to the zinc- and cobalt-bound vCP1 (*cf*. panels ***B*** and ***C*** *vs* ***D*** and ***E*** in ***Figure 3-2***). This is indirect evidence that Zn^{2+} and Co^{2+} are specifically bound within the peptide's coordination sphere, which is not the case for Ca^{2+} and Cu^{2+}.

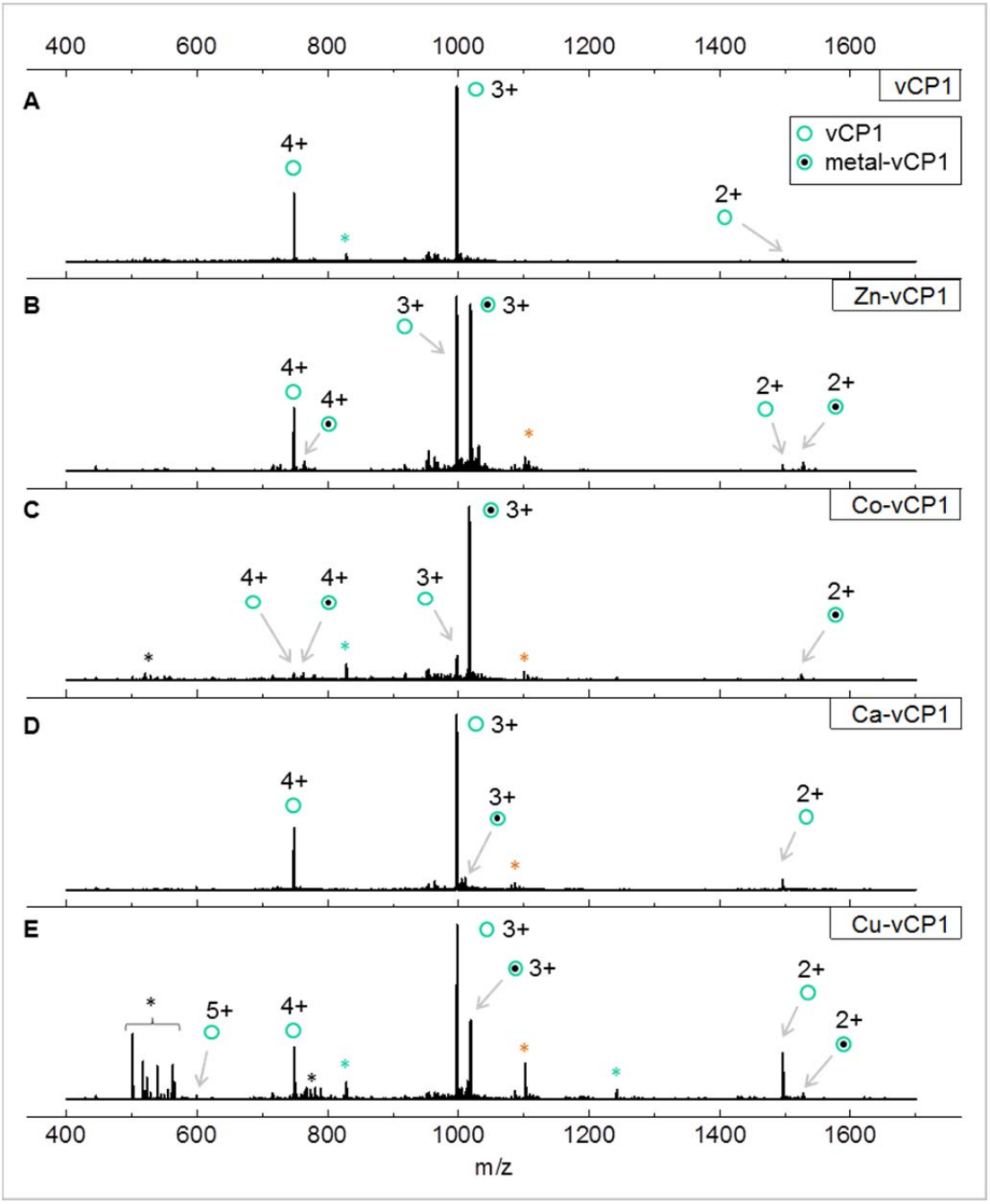


Figure 3-2 *Representative nESI mass spectra of vCP1 individually and with metals. (**A**) – vCP1 sprayed from the buffered conditions (see below); (**B**) through (**E**) – equimolar mixture of vCP1 and metal acetate salts, featuring apo and holo forms of the peptide: (**B**) – Zn, (**C**) – Co, (**D**) – Ca, (**E**) – Cu. Conditions: 20 μM peptides and metal acetate, 10 mM ammonium acetate, 5% isopropanol, 200 μM TCEP, pH 6.8. The peaks are denoted as follows (see legend in pane **A**: one for all spectra): apo-vCP1 by turquoise circle and holo-vCP1 by turquoise circles with a black dot in the centre. Impurities resulting from metal salts are shown in black asterisks; orange asterisks show TCEP adducts on the 3+ charge state ions. In-source fragmentation of the peptide (at its proline-5 site) is shown in turquoise asterisks.*

To be able to explain the observed differences in relative peak intensities for different metal ions one needs to investigate the events taking place in the metal coordination sphere. Here, a crucial role in binding of metal ions is played by the oxidation states of the two Cys amino acid residues. Isotopic cluster analysis, detailed in the following section, will help shed light on these events.

3.4 Metal ions and vCP1 – quantifying metal ion affinity

Figure 3-3 shows typical nESI mass spectra for vCP1 without (***Figure 3-3A***) and with (***Figure 3-3B*** though ***E***) equimolar amounts of metal salts added; the regions for the 3+ charge state(s) of the peptide are shown. The $[M + 3H]^{3+}$ or equally charged $[M + X + H]^{3+}$, where X = metal in 2+ oxidation state, were the dominant peaks in the spectra, ~8 times more intense than the 2+ and 4+ species. Use of TCEP maintained approximately 99% of the cysteines in the fully reduced state (sulphur oxidation state 2–) as evidenced by isotopic cluster analysis (see insets in ***Figure 3-3***). This is not the case for the Cu^{2+}, which will be discussed later in this section. Theoretical fitting for the elemental compositions of the fully reduced peptide is superimposed on the experimental isotopic distribution, which leads to conclusion that under these conditions the metals are coordinated by thiolates ($–S^-$), sulphur oxidation state 1–, and the metals are in the 2+ oxidation state[28]. Impurity peaks annotated on ***Figure 3-3*** are the common adducts of oxygen (*) and sodium (**), as well as chloride (***) from TCEP salt. The ratio between the intensity of the ^{12}C peak for both apo and holo species was determined for

each spectrum, which estimates an order of metal affinity for vCP1 as Co^{2+} > Zn^{2+} > Cu^{2+} >> Ca^{2+}.

Preferences for 1:1 vCP1:metal binding were observed for Co^{2+} (***Figure 3-3C***), and to a slightly lesser extent for Zn^{2+} (***Figure 3-3B***) compared to the other metal ions. For both of these metals the same set of adducts, and at similar relative intensities as observed for apo-vCP1, were present with the holo-forms. Perhaps not surprisingly, in the presence of Cu^{2+} there was a higher fraction of the oxidised form of apo-vCP1, as revealed by the isotopic cluster analysis (***Figure 3-3E***, left inset), despite the 50-fold excess of TCEP. This is evidenced by the two fronting low-intensity peaks featured on the isotopic cluster: these form part of the apo-vCP1 sub-population that is oxidised and therefore has a lower mass. For Cu-vCP1 (***Figure 3-3E***, right inset), the red dots represent a theoretical fit for a similar scenario as for the other metals, *i.e.* that both cysteines are in a thiolate form ($–S^-$). In reality, the experimental data fit when only one cysteine is in the thiolate form, and the other one is in thiol form (–SH) as denoted by red stars. Hence, on interaction with sulphydryl groups of cysteines, Cu^{2+} ions were reduced to Cu^{1+}. It has been previously reported that copper can catalyse the formation of disulphide bridges between two Cys[29] leading to oxidative modification of proteins[30].

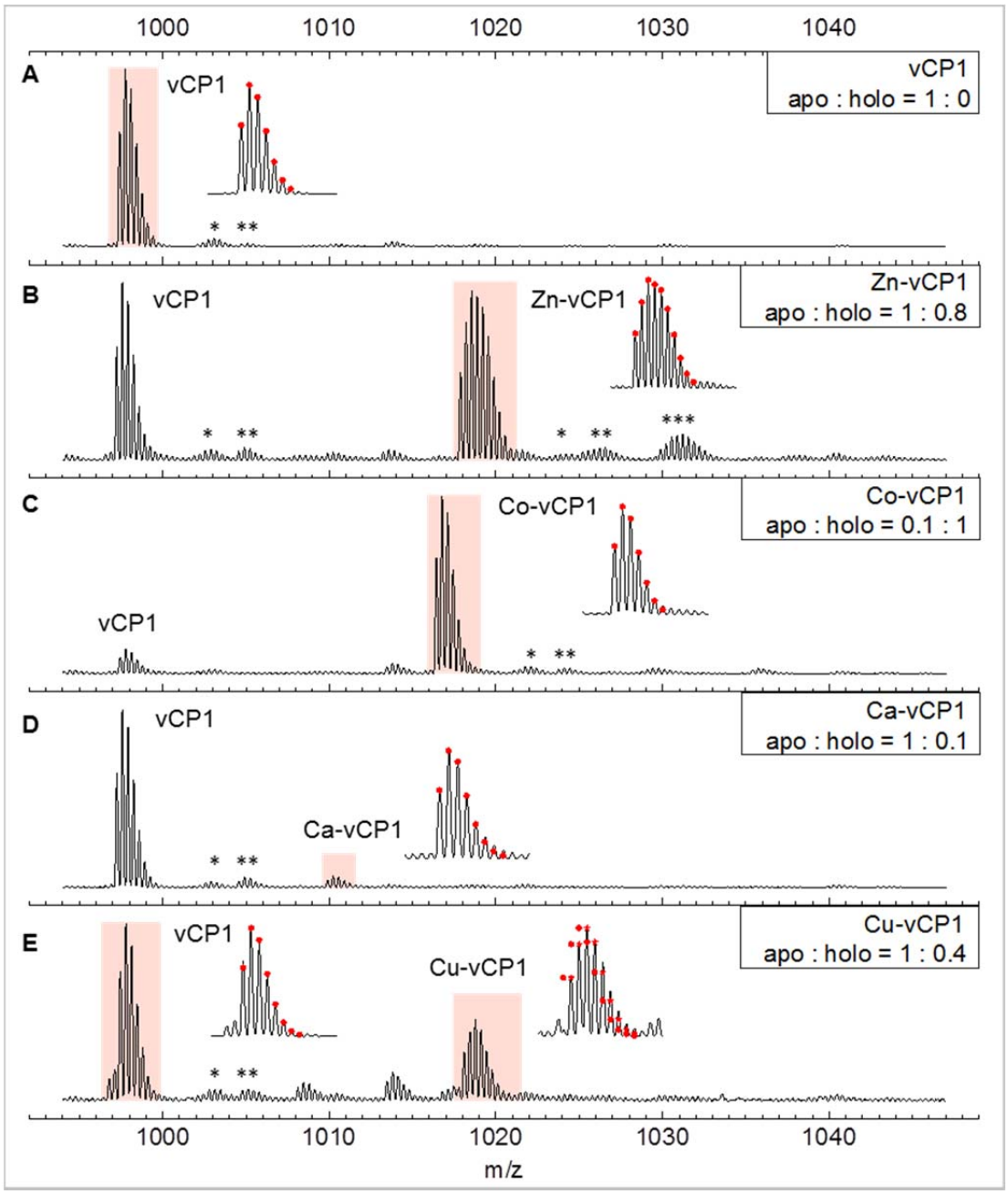


Figure 3-3 *Representative nESI mass spectra of the 3+ charge state of vCP1 for: (**A**) the apo state, and with: (**B**) Zn^{2+}, (**C**) Co^{2+}, (**D**) Ca^{2+} and (**E**) Cu^{2+}. Ratios of apo to holo form were calculated from the intensities of the mono-isotopic peaks. Insets show the resolved isotopic clusters for the free and bound states (highlighted in pink on the mass spectra) along with theoretical fitting (red dots). Isotopic cluster analysis indicated that 99% of the peptide is reduced, except in the case of Cu^{2+} (see main text). Impurity peaks are annotated with asterisks: common MS adducts of oxygen (*) and sodium (**); chloride adduct originating from TCEP (***). Conditions: 20 μM peptide; 10 mM ammonium acetate; 5% isopropanol; pH 6.8; 200 μM TCEP (1 mM TCEP used with CuAc); 20 μM metal acetate salts.*

Results presented here show that calcium binds to vCP1 with lowest affinity in both the gas and solution phase (***Figure 3-3D*** and ***Figure A 3*** in ***Appendix***). A small amount of calcium (a contaminant of lab-ware, reagents, and even deionised water) was found to bind the peptide even in the presence of Zn^{2+}, leading to the conclusion that Ca^{2+} binds vCP1 non-specifically, and probably away from the Cys_2His_2 binding site (this supports the hypothesis made in section ***3.3*** that the most likely site for Ca^{2+} to bind on vCP1 is the Glu-6 residue).

Gas-phase dissociation constants K_d were determined for all metals interacting with vCP1 following the procedure outlined in section ***2.6.1***. In order to determine K_d for metal binding all charge states were considered. To determine the equilibrium constants for the vCP1 binding metal ions, the peptide was incubated with increasing concentrations of metal acetate salt at pH 6.8, and the first monoisotopic signal intensities for the apo and holo-vCP1 were summed for all charge states. Typically, the following assumptions are made when calculating K_ds from ESI-MS measurements[31; 32; 33]:

❖ The mass and size of the ligand (metal ion in the case here) is small relative to the target peptide, so its contribution (negative or positive) to the overall ionisation efficiency is negligible when comparing free and bound peptide.

❖ Peak intensities from mass spectra correlate to concentrations in solution at the relatively low concentrations that are used for nESI-MS, as this method preserves non-covalent interactions in complexes, when used correctly. Consequently, the relative ratios of bound *versus* non-bound target

are considered to reflect of the ratios in solution. This latter point depends on the interaction strength, but is likely to hold for electrostatic interactions as is the case here.

In this study, the intensity of the first monoisotopic peak is chosen as a measure of concentration, due to its ease of implementation. Preliminary experiments (data not shown) demonstrated that this method yielded comparable results to the more laborious integration of the area under the monoisotopic peak. Due to natural isotopic distribution, the method employed here has a bias towards under-estimating binding of zinc (by 51%), calcium (by 3%) and copper (by 31%). Although numerically these values may be significant, the order of vCP1 affinity towards metals remains the same. Relative intensities of the first monoisotopic MS signals have been recalculated into concentration units, the data from three experimental repeats have been averaged, and the ratio of molar concentration of bound metal ion to total molar concentration of the vCP1 was plotted against the concentration of the free ligand. A Hill function has been fitted to the data points, and the K_d values obtained at 50% system saturation, *i.e.* half of the peptide population binds metal. The Hill binding model is generally used for describing cooperative binding when more than one binding site is present on the macromolecule[34; 35]. The same relationship that describes a sigmoidal curve can also be used in cases of non-cooperativity (a special case[34; 35] when only one biding site is present), in which case it becomes a classical Michaelis-Menten expression whose plot is hyperbolic. If a given peptide or protein had only one binding site, then cooperative behaviour would not be expected, and in such cases the sigmoidal signal response curve would be assigned to a switching phenomenon[36]. However, in order to properly test protein design the analytical method must be impartial; therefore in this

thesis a Hill function is employed to examine binding, since it will describe both cooperative and non-cooperative behaviour.

Figure 3-4A through ***D*** shows binding curves obtained for vCP1 complexed by Zn, Co, Ca and Cu, respectively. In good agreement with the results shown in ***Figure 3-3***, strong binding is observed for Co^{2+} and Zn^{2+} (with K_ds of 0.15 ± 0.01 and 12 ± 0.1 μM respectively), poorer for Cu^{2+} (K_d = 88 ± 28 μM) and very poor for Ca^{2+} (K_d = 483 ± 24 μM) where no full binding has been achieved. However, at low concentrations, cobalt appears to give stronger binding than zinc (*cf*. ***Figure 3-4A*** and ***B***). This agrees with the observation made earlier that vCP1 binds cobalt at low concentrations with greater affinity than zinc. The low nanomolar range of K_d values for cobalt binding to vCP1 is on the limit of detection for such measurements performed in the gas phase. The resulting value of K_d is therefore innately prone to a large margin of error. The sigmoidal shape of the zinc-binding curve, as well as the value of $n > 1$, points to the switching behaviour of the vCP1 when presented with Zn^{2+} cation. Considering the physical similarities of Zn^{2+} and Co^{2+}, it would probably be reasonable to suggest that the cobalt-binding curve also is sigmoidal at very low metal concentrations, with a short 'lag phase' that is not detectable by this method. Also, it must be pointed out that MS-based method for measuring binding affinity has its lower applicability limit at the K_d values of ~2 μM. Therefore, any value measured below that is innately inaccurate and should be interpreted as 'tight binding with K_d values below 2 μM', as is the case with cobalt binding (***Figure 3-4B***).

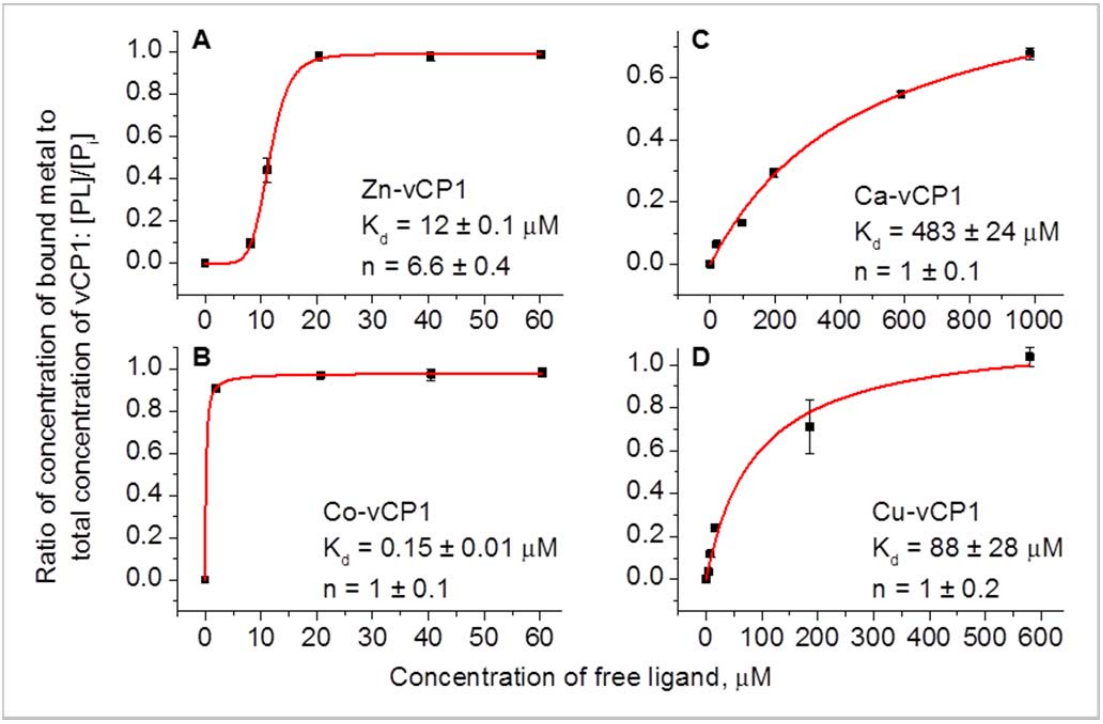


Figure 3-4 *Titration curves of vCP1 against metals obtained from titration nESI-MS experiments: (**A**) Zn^{2+}, (**B**) Co^{2+}, (**C**) Ca^{2+}, (**D**) Cu^{2+}. The K_ds are determined by keeping the vCP1 concentration at 20 μM and titrating the metals in the following concentration ranges: Zn and Co from 0 to 80 μM; Ca from 0 to 1 mM; Cu from 0 to 600 μM. Hill functions are fitted to the data points and the K_d values obtained at 50% system saturation, i.e. half of the peptide population binds metal. The K_d values and number of cooperative sites n is quoted on each panel. Data are derived from relative ion currents in ESI mass spectra (first monoisotopic peak), summing the intensities of ion currents for all charge states of each species. Each data point is the mean (± standard error) of the equilibrium concentrations from three mass spectra.*

At the ligand concentration range considered here, non-saturable binding is observed for Ca^{2+} (***Figure 3-4C***); this is indicative of non-specific nature of it, which is evident from the calcium binding curve that reaches only ~0.7 saturation. Calcium is also the only metal out of the four studied here that does not have a two-metal binding peak on the mass spectrum. These findings corroborate the point made in section ***3.3***, that calcium binds to carboxyl group of the Glu-6 residue – the only most probable site available for non-specific binding. Copper binding curve (***Figure 3-4D***) is saturated, suggesting specificity of the metal binding in the Cys_2His_2 coordination sphere, albeit in a thiol-thiolate arrangement of the two cysteines, which is

different from how Zn^{2+} and Co^{2+} are bound (thiolate-thiolate). The curve shapes for both Ca^{2+} and Cu^{2+} are hyperbolic, indicating that no conformational switching takes place as a result of metal binding.

Considering this, one might conclude that, there is a good agreement with the UV-visible spectroscopic data that estimated the K_d for cobalt to be 0.8 μM (carried out by the collaborator, see ***Appendix***, ***Figure A 4***). The spectroscopic data for zinc binding estimated its K_d to be sub-nanomolar and therefore very tight. Such a vast difference between binding strength of the two very similar metals may be attributed to the experimental setup whereby the zinc dissociation constant is measured by displacement of cobalt from its coordination sphere. In this case, the structure of binding site remains 'configured' for metal binding, as it does not relax quickly enough upon cobalt leaving, resulting in apparent observation of tighter zinc binding. Such displacement experiment has not been carried out in the gas phase for the vCP1 system, however this is implemented for the similar zinc finger system discussed in the following chapter (with the results supporting the hypothesis of 'binding site configuration'). Another point, of a more general character, that needs to be highlighted here is that direct comparisons between solution- and gas-phase studies of non-covalent interactions should be done with a degree of caution. Electrospray conditions need to be controlled carefully in order to preserve such interactions, and even in this case they will be weaker than those in solution[37]. Also, the relative contribution of inter- and intra-molecular forces to preservation of the protein complex may differ between solution and gas phase[38]. Although it is very important not to over-interpret the results of such comparative studies, the gas-phase method can provide a useful insight into the nature of protein-ligand interactions.

3.5 Collision cross sections and molecular dynamics simulations – elucidating conformations

By performing IM-MS experiments for all of the apo and the holo forms of vCP1 shown in ***Figure 3-3***, arrival time distributions (ATDs) were obtained at a range of drift voltages which are converted to collision cross sections following ***Equation 2-5*** in ***Chapter 2***, and the CCS values of the 3+ charge state are presented in ***Table 3-1***. The CCS values of apo-vCP1 measured from the spray solution not containing any metal salts are listed in the top row of the ***Table 3-1***. All values for the apo-vCP1 below that are obtained from the unbound form of the peptide present in solution containing metal salts to the concentration necessary to observe both apo and holo forms. This was done to investigate whether the metal ion presence affects apo-vCP1 conformation. Findings here are inconclusive, as the relevant CCS values of apo and holo forms presented in ***Table 3-1*** are similar and within the margin of error of each other. The packing found for the apo-vCP1 and Zn-vCP1 gave the lowest, and surprisingly similar, values. The expectation for the CCS values upon zinc and cobalt binding was to be consistent with K_d findings, that is to observe conformational tightening effected by specific metal binding (in contrast to calcium and copper, that do not bind specifically). Although the data presented in ***Table 3-1*** show no indication of CCS difference between specific and non-specific binding, the results indicate that overall packing of the vCP1 tends to be independent of the metal binding, and the difference may lie only in the arrangement of the metal-coordinating residues.

	Apo-vCP1 (Å^2)	Holo-vCP1 (Å^2)
vCP1	488 ± 13	-
Zn-vCP1	487 ± 14	488 ± 12
Co-vCP1	518 ± 2	492 ± 10
Cu-vCP1	499 ± 17	493 ± 13
Ca-vCP1	528 ± 18	517 ± 12

Table 3-1 *Experimental collision cross sections of the most abundant charge state (3+) of the apo and holo-vCP1 (four metals) in 'buffered' conditions. The values are quoted in Å^2, with standard error of three experimental repeats.*

To investigate this hypothesis further, molecular dynamics (MD) simulations of vCP1 were performed on its apo and zinc-bound form. ***Figure 3-5*** shows representative nESI mass spectra and ATD data for the 3+ charge state of apo-vCP1 and Zn-vCP1, along with structures obtained by MD simulations[39]. Experimental collision cross sections of the apo-vCP1 and Zn-vCP1 compare favourably with those obtained from MD giving confidence in both datasets. Interestingly, the apo form appears to possess some helical character (***Figure 3-5***), although we find these helical regions are transient (appearing at different regions of the sequence along the MD trajectory), particularly when compared to the holo form. A substantial difference is observed in the coordination sphere: the side chains of Cys and His residues are further apart in the absence of the metal ion. When solvent is present in the simulations, however, there is little evidence for any helicity in the apo peptide (not shown).

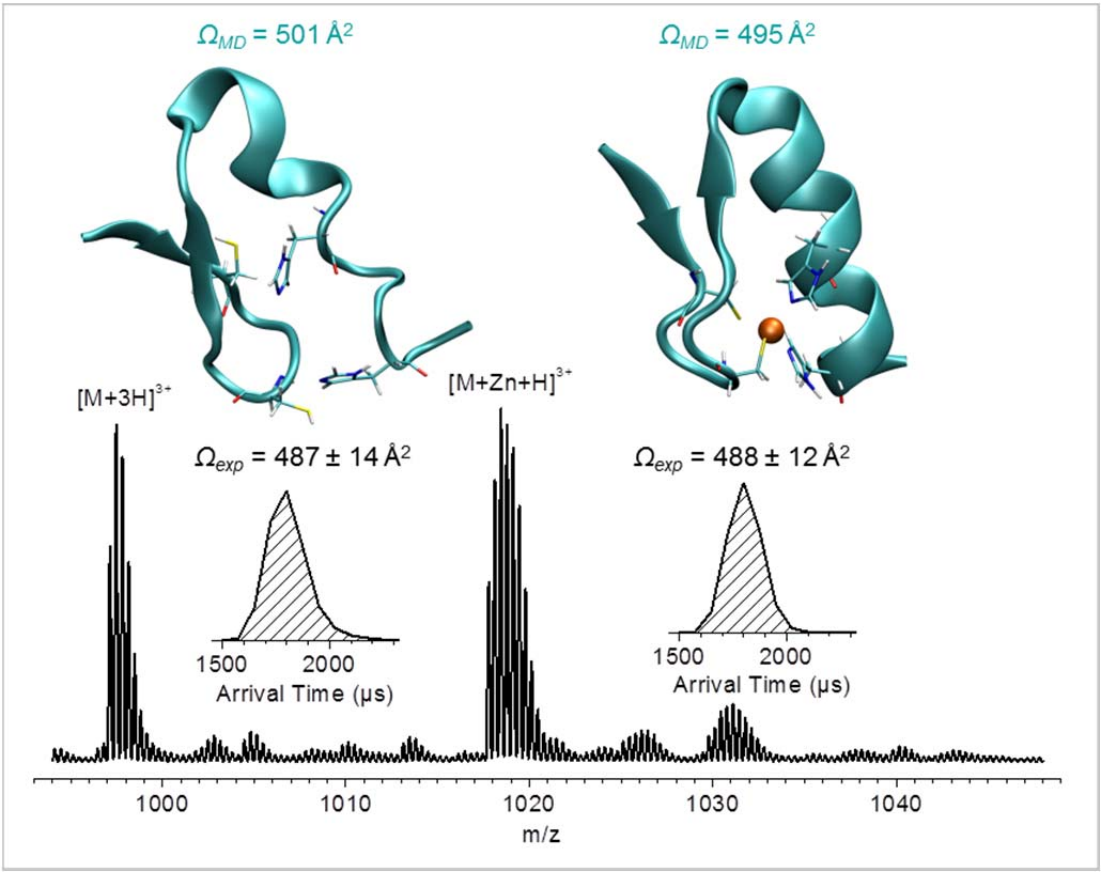


Figure 3-5 *nESI mass spectra of the 3+ charge state obtained by spraying the 1:1 ZnAc:vCP1 from same buffer conditions as in* ***Figure 3-3****. Insets show the ATDs of apo and holo-vCP1, and representative MD structures, along with collision cross sections obtained experimentally and from the MD simulations.*

3.6 Conclusions

It has been illustrated how gas-phase measurements can complement those in solution to provide information on the binding stoichiometry, specificity and affinity of a model metal-binding peptide system. As a label-free platform, the gas-phase interrogation approach offers an attractive alternative to solution-phase methods to assist protein folding studies and inform protein design efforts. The primary findings for vCP1 system are:

❖ Qualitatively the agreement between the methods is excellent, and some semi-quantitative understanding can be gained also.

❖ By MS, there is strong evidence that vCP1 preferentially and specifically binds Zn^{2+} and Co^{2+}, when compared to Ca^{2+} and Cu^{2+}. Analysis of charge state distributions suggests increased conformational stability upon binding Zn^{2+} and Co^{2+} by vCP1.

❖ By CD, in the absence of metal ions the peptide is largely unstructured, whereas upon the addition of zinc or cobalt the peptide folds to give a spectrum consistent with that reported for naturally occurring zinc fingers. In contrast, addition of copper and calcium changes the conformation of the peptide in a less pronounced manner.

❖ UV absorbance spectroscopy (using Co^{2+} as a spectroscopic probe) revealed tetrahedral coordination of the metal by vCP1.

❖ The use of IM-MS provides further insight into coordination behaviour as well as the conformational preferences of a zinc finger peptide. These results corroborate findings of MD studies on the peptide.

❖ Using vCP1, it has been demonstrated that ion mobility mass spectrometry is a robust tool for protein design: its findings can inform design strategy.

3.7 References

1. James, L. C. & Tawfik, D. S. (2003). Conformational diversity and protein evolution - a 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **28**, 361-368.

2. Ramos, C. H. I. & Ferreira, S. T. (2005). Protein folding, misfolding and aggregation: evolving concepts and conformational diseases. *Protein Pept. Lett.* **12**, 213-222.

3. Lee, D., Redfern, O. & Orengo, C. (2007). Predicting protein function from sequence and structure. *Nat. Rev. Mol. Cell Biol.* **8**, 995-1005.

4. Bromley, E. H. C., Channon, K., Moutevelis, E. & Woolfson, D. N. (2008). Peptide and protein building blocks for synthetic biology: From programming biomolecules to self-organized biomolecular systems. *ACS Chem. Biol.* **3**, 38-50.

5. Channon, K., Bromley, E. H. C. & Woolfson, D. N. (2008). Synthetic biology through biomolecular design and engineering. *Curr. Opin. Struct. Biol.* **18**, 491-498.

6. Gunasekar, S. K., Haghpanah, J. S. & Montclare, J. K. (2008). Assembly of bioinspired helical protein fibers. *Polym. Adv. Technol.* **19**, 454-468.

7. Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L. & Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364-1368.

8. Li, W., Zhang, J. & Wang, W. (2007). Understanding the folding and stability of a zinc finger-based full sequence design protein with replica exchange molecular dynamics simulations. *Proteins Struct. Funct. Bioinf.* **67**, 338-349.

9. Shi, Y., Beger, R. D. & Berg, J. M. (1993). Metal binding properties of single amino acid deletion mutants of zinc finger peptides: studies using cobalt(II) as a spectroscopic probe. *Biophys. J.* **64**, 749-753.

10. Wilson, G. S. & Hu, Y. (2000). Enzyme-Based Biosensors for in Vivo Measurements. *Chem. Rev. (Washington, D. C.)* **100**, 2693-2704.

11. Smith, R. G., D'Souza, N. & Nicklin, S. (2008). A review of biosensors and biologically-inspired systems for explosives detection. *Analyst (Cambridge, U. K.)* **133**, 571-584.

12. Chockalingam, K., Blenner, M. & Banta, S. (2007). Design and application of stimulus-responsive peptide systems. *Protein Eng. Des. Sel.* **20**, 155-161.

13. Yue, K. & Dill, K. A. (1992). Inverse protein folding problem - designing polymer sequences. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4163-4167.

14. Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E. & Thornton, J. M. (1987). Knowledge-based prediction of protein structures and the design of novel molecules. *Nature* **326**, 347-352.

15. Branden, C. & Tooze, J. (1992). *Introduction to Protein Structure*, Kyoikusha Ltd.

16. Leatherbarrow, R. J. & Fersht, A. R. (1986). Protein engineering. *Protein Eng.* **1**, 7-16.

17. Berl, V., Huc, I., Khoury, R. G., Krische, M. J. & Lehn, J.-M. (2000). Interconversion of single and double helices formed from synthetic molecular strands. *Nature (London)* **407**, 720-723.

18. Delsuc, N., Godde, F., Kauffmann, B., Leger, J.-M. & Huc, I. (2007). The Herringbone Helix: A Noncanonical Folding in Aromatic-Aliphatic Peptides. *J. Am. Chem. Soc.* **129**, 11348-11349.

19. Petitjean, A., Cuccia, L. A., Schmutz, M. & Lehn, J.-M. (2008). Naphthyridine-Based Helical Foldamers and Macrocycles: Synthesis, Cation Binding, and Supramolecular Assemblies. *J. Org. Chem.* **73**, 2481-2495.

20. Bellows-Peterson, M. L., Fung, H. K., Floudas, C. A., Kieslich, C. A., Zhang, L., Morikis, D., Wareham, K. J., Monk, P. N., Hawksworth, O. A. & Woodruff, T. M. (2012). De Novo Peptide Design with C3a Receptor Agonist and Antagonist Activities: Theoretical Predictions and Experimental Validation. *J. Med. Chem.* **55**, 4159-4168.

21. Kim, C. A. & Berg, J. M. (1996). A 2.2 A resolution crystal structure of a designed zinc finger protein bound to DNA. *Nat. Struct. Mol. Biol.* **3**, 940-945.

22. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L. & Berg, J. M. (1991). A consensus zinc finger peptide: design, high-affinity metal binding, a pH-dependent structure, and a His to Cys sequence variant. *J. Am. Chem. Soc.* **113**, 4518-4523.

23. Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T. & Berg, J. M. (1992). Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4796-800.

24. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A. & Wright, P. E. (1989). Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science (Washington, D. C., 1883-)* **245**, 635-7.

25. Berg, J. M. (1990). Zinc Finger Domains: Hypotheses and Current Knowledge. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 405-421.

26. Litovchick, A. & Szostak, J. W. (2008). Selection of cyclic peptide aptamers to HCV IRES RNA using mRNA display. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15293-15298.

27. Kretsinger, R. H. (1976). Calcium-binding proteins. *Annu. Rev. Biochem.* **45**, 239-66.

28. Lei, Q. P., Cui, X., Kurtz, D. M., Amster, I. J., Chernushevich, I. V. & Standing, K. G. (1998). Electrospray Mass Spectrometry Studies of Non-Heme Iron-Containing Proteins. *Anal. Chem.* **70**, 1838-1846.

29. Petros, A. K., Reddi, A. R., Kennedy, M. L., Hyslop, A. G. & Gibney, B. R. (2006). Femtomolar Zn(II) Affinity in a Peptide-Based Ligand

Designed To Model Thiolate-Rich Metalloprotein Active Sites. *Inorg. Chem.* **45**, 9941-9958.

30. Stadtman, E. R. (1990). Metal ion-catalyzed oxidation of proteins: Biochemical mechanism and biological consequences. *Free Radical Biol. Med.* **9**, 315-325.

31. Gabelica, V., Galic, N., Rosu, F., Houssier, C. & De Pauw, E. (2003). Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. *J. Mass Spectrom.* **38**, 491-501.

32. Boeri Erba, E., Barylyuk, K., Yang, Y. & Zenobi, R. (2011). Quantifying Protein–Protein Interactions Within Noncovalent Complexes Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **83**, 9251-9259.

33. Sannes-Lowery, K. A., Griffey, R. H. & Hofstadler, S. A. (2000). Measuring Dissociation Constants of RNA and Aminoglycoside Antibiotics by Electrospray Ionization Mass Spectrometry. *Anal. Biochem.* **280**, 264-271.

34. Monod, J., Wyman, J. & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88-118.

35. van Holde, K. E., Johnson, W. C. & Ho, P. S. (1998). *Principles of Physical Biochemistry*, Prentice-Hall, Inc.

36. Goldbeter, A. & Koshland, D. E. (1981). An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6840-6844.

37. Robinson, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M. & Dobson, C. M. (1996). Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein - CoA Ligand Binding and Assembly. *J. Am. Chem. Soc.* **118**, 8646-8653.

38. Breuker, K. & McLafferty, F. W. (2008). Stepwise evolution of protein native structure with electrospray into the gas phase, 10E-12 to 10E2 s. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18145-18152.

39. Shvartsburg, A. A., Schatz, G. C. & Jarrold, M. F. (1998). Mobilities of carbon cluster ions: Critical importance of the molecular attractive potential. *J. Chem. Phys.* **108**, 2416-2423.

Chapter 4

Ion mobility mass spectrometry as a tool for protein design. Part 2: a case study on zinc finger fold *versus* coiled coil interactions

4.1 Introduction

Presented in this chapter is development of a gas-phase analytical platform for interrogation of a peptide-based dual switching system (ZiCop) that is designed to adopt either of the two conformations in response to a stimulus: zinc finger or coiled coil. Some elements of the characterisation method have been test-run using a consensus zinc finger peptide vCP1, and results presented in the previous chapter. The zinc-bound state of the switch is very similar to the one described using the example of vCP1. The coiled coil state involves the peptide's interaction with another peptide to form a stable two-helix parallel dimeric coiled coil.

Some aspects of understanding of sequence-to-structure relationships in proteins, as well as some practical steps in approaching protein design along with perils encountered by synthetic biologists, have been outlined in the previous chapter. At this point, a few introductory words need to be said about incorporating a conformational switching functionality into a designed system.

Protein design is not a straightforward process, as elucidating an amino acid sequence that is compatible with a certain conformation does not necessarily lead to that fold (more often it does not)[1; 2]. This is not surprising, as in Nature many protein functions are underpinned by a conformational change affected by a stimulus. Examples of this are numerous, to name just a few: lymphotactin that plays an important role in the immune system[3], adenylate kinase that maintains energy balance in the cell[4], calmodulin that mediates many cell regulation processes[5]. Along with protein function being dependent on conformational changes, paradoxically, protein malfunction is often caused by these changes too. Notoriously devastating diseases such as Alzheimer's and Parkinson's have conformational transitions as the key events in their onset[6].

Here, gas-phase methodologies are used to monitor such conformational transitions using a synthetic dual switching system that is designed to adopt one of two folds in response to stimulus. Presented in this chapter is the use of mass spectrometry (MS) and ion mobility mass spectrometry (IM-MS) coupled with molecular dynamics (MD) simulations as an analytical platform to inform *de novo* design efforts for peptide-metal and peptide-peptide interactions. A synthetic dual peptide-based system, ZiCop, based on a zinc finger peptide motif has been investigated. Titration mass

spectrometry determines the relative binding affinities of different divalent metal ions[7] as $Zn^{2+} > Co^{2+} >> Ca^{2+}$. With collision induced dissociation (CID), complex stability is probed to establish that peptide-metal interactions are stronger and more 'specific' than those of peptide-peptide complexes, and the anticipated hetero-dimeric complex is more stable than the two homo-dimers. Collision cross sections (CCS) measurements by IM-MS reveal increased stability with respect to unfolding of the metal-bound peptide over its apo-form, and further, larger collision cross sections for the hetero-dimeric forms suggest that dimeric species formed in the absence of metal are coiled coil like. MD supports these structural assignments which are also backed up by data from visible light absorbance and circular dichroism measurements.

Determining the CCS provides conformations adopted in the gas phase[8; 9] which can be related to solution conformations. Experimentally derived CCS and hence conformations, can be complemented by MD analysis to atomistically resolve the structures[10; 11; 12]. IM-MS has been successfully applied in studies of sequence-to-structure relationships in polypeptides, since structures observed *in vacuo* are defined intrinsically by the sequence[7; 13; 14].

4.2 Design of ZiCop switching system

The peptide presented in this work is a dual 'new generation' of a switch designed to be capable of reversible transition between zinc finger and coiled coil conformations. An earlier attempt to design a conformational switch has been reported by the Woolfson group[15] as discussed in section ***1.1.2.6***. The

dual switching system discussed in this chapter is a peptide that encodes two different sequences for two distinct folds simultaneously; the transition between conformational states is reversible and is triggered by specific metal binding. Presented here is a mass spectrometry based analytical workflow that can be used to characterise and test *de novo* designed peptides (***Figure 4-1***). Herein, a characterisation platform has been developed to interrogate the following parameters: stoichiometry of binding, metal affinity, fold, conformational transitions, and the strength of interactions between hetero-dimers. Mass spectrometry (MS), ion mobility mass spectrometry (IM-MS), molecular dynamics (MD) and solution based measurements (the latter carried out by the collaborator on this project at the University of Bristol) are used in the current work.

The two conformations of the test switch are zinc finger and coiled coil folds. The reversible switch between the two conformational states is controlled by the binding of a metal ion, zinc, to the peptide (***Figure 4-1A***). The aim of synthesis was that the peptide, named ZiCop (reflecting the structural duality of zinc finger and coiled coil peptide), in the absence of the Zn^{2+} ions would form and amphipathic parallel dimeric helix for coiled coil interactions with a partner peptide (Pp), then with the addition of Zn^{2+} form a metal-bound monomer adopting a zinc finger conformation. The system can be switched back to the coiled coil interactions through the removal of Zn^{2+} using EDTA.

In early versions of the peptide it was found that zinc-ligating cysteine residues had a tendency to oxidise[15], therefore a reducing agent (TCEP) was added to the system to prevent this. ZiCop was designed by merging alignments of coiled coil and zinc finger sequences. These were then used to

generate sequences that were characteristic of both folds. These sequences were then filtered according to their potential to form parallel hetero-dimeric coiled coils with designed partner peptides. The partner peptide was designed through analysing experimentally determined side-chain interaction energies found in parallel dimeric coiled coils[16; 17].

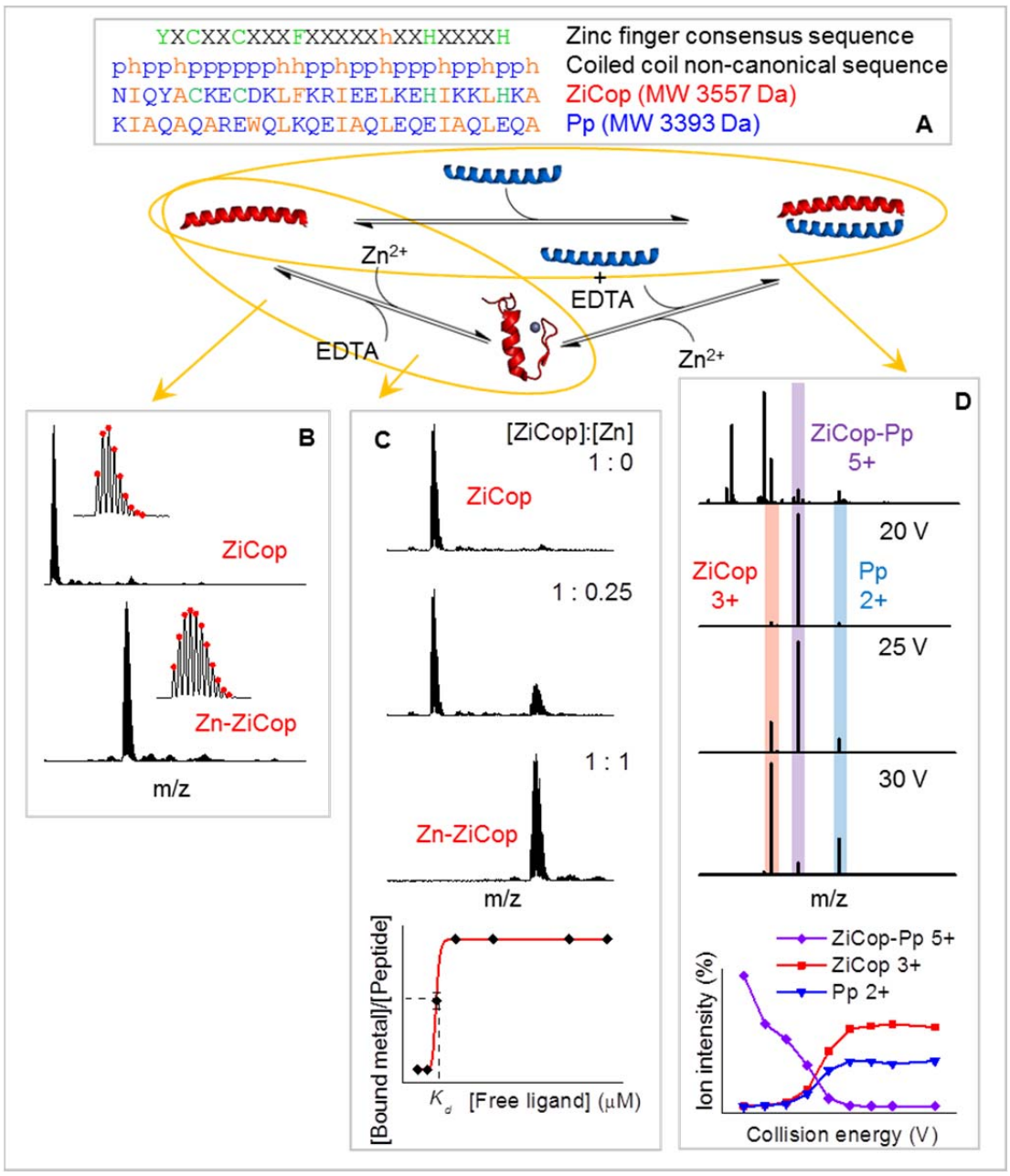


Figure 4-1 *Design concept of the dual reversible peptide-based switch. (**A**) – Sequence of the ZiCop and its binding partner Pp in relation to the structural requirements for the zinc finger and coiled coil designs. The 'X' refers to any amino acid, 'h' – to any hydrophobic residue, 'p' – to any polar residue. Below is proposed switching process: the red peptide represents ZiCop, which is responsive to zinc; the partner peptide (blue) was designed to interact with ZiCop in the absence of zinc. (**B**) through (**D**) – A schematic representing analytical workflow of the design concept interrogation using MS platform. (**B**) – Stoichiometry of binding (top spectrum: free ZiCop peak, bottom: Zn-bound peptide). (**C**) – Quantification of affinity of the peptide to the metal ion (mass spectra of Zn titrations into ZiCop and resulting biding curve to yield dissociation constant). (**D**) – Evaluation of the hetero-dimeric coiled coil association using collisionally induced dissociation (CID); resulting curves indicate survival of the precursor ion of the coiled coil (purple) and formation of the products – ZiCop and Pp (red and blue respectively).*

Both resulting peptides have sequences of 29 amino acids in length (***Figure 4-1A***), and the schematic featuring proposed binding behaviour of this system is shown below the sequences. In the absence of the Zn^{2+} ion, a parallel hetero-dimeric coiled coil is formed by the two peptides. When the metal is added, one of the interacting partners – ZiCop, should form a zinc finger fold around the metal. Mass spectrometry can monitor this process by detecting the mass shift of the molecular ion of ZiCop upon metal binding (***Figure 4-1B***). A quantitative insight into the metal-binding event can be obtained by titrating increasing amounts of the metal into the peptide and measuring relative change in the MS peak abundance of the apo- and holo-forms of the ZiCop (***Figure 4-1C***). By converting peak intensity into concentration units, a binding curve can be plotted, from which a dissociation constant K_d can be obtained. ***Figure 4-1D*** demonstrates how the binding strength of the hetero-dimeric coiled coil can be explored by collision induced dissociation of the complex. Mass spectra of the precursor ion and resulting products are recorded as a function of collision voltage and the intensities of all the channels are converted to concentration units yielding plots of the precursor survival and product formation as a function of the voltage applied.

In this study IM-MS coupled with MD is used to determine the conformations of peptides and their complexes, and the use of this technique is explored as a tool to assist protein design efforts to encode fold from primary sequence.

4.3 Overview of interactions – stoichiometry of binding

Under the experimental conditions described in ***Chapter 2***, ***Table 2-1***, the ZiCop peptide presents two major charge species assigned as $[ZiCop+3H]^{3+}$ and $[ZiCop+4H]^{4+}$ (***Figure 4-2A***), with the latter being the most dominant, the former is approximately half its intensity. The Partner peptide Pp also presents these charge states, but here $[Pp+3H]^{3+}$ is more dominant (***Figure 4-2B***). In both cases, dimerisation occurs to yield a homo-dimer predominantly as a $[2M+5H]^{5+}$ ion. Isotopic cluster analysis suggests that a smaller amount of the $[2M+4H]^{4+}$ ion is also present, but this is *m*/*z*-coincident with low abundant $[M+2H]^{2+}$ monomer. When the two peptides are mixed together in the equimolar amounts (***Figure 4-2C***), a small peak at *m*/*z* 1391.0 is observed corresponding to the hetero-dimer $[Pp:ZiCop+5H]^{5+}$. The relatively low intensity of this species, compared with that of the constituent monomers suggests that the interactions between the two peptides are not particularly strong. Addition of zinc to the peptides, to give equimolar amounts of each component (***Figure 4-2D***), yields full binding of the ZiCop peptide to the metal, accompanied by the loss of the $[ZiCop+5H]^{5+}$ species (*cf.* ***Figure 4-2C*** and ***D***), suggesting a more compact conformation with fewer protonatable groups accessible to solvent. In addition, in the presence of zinc, the (now metal bound) monomeric ZiCop species drops in intensity relative to the Pp. Interestingly, the zinc-bound form of the ZiCop tends to form a dimer and also to interact with the Pp, with the latter effect observed by spectroscopic analysis also (data not shown).

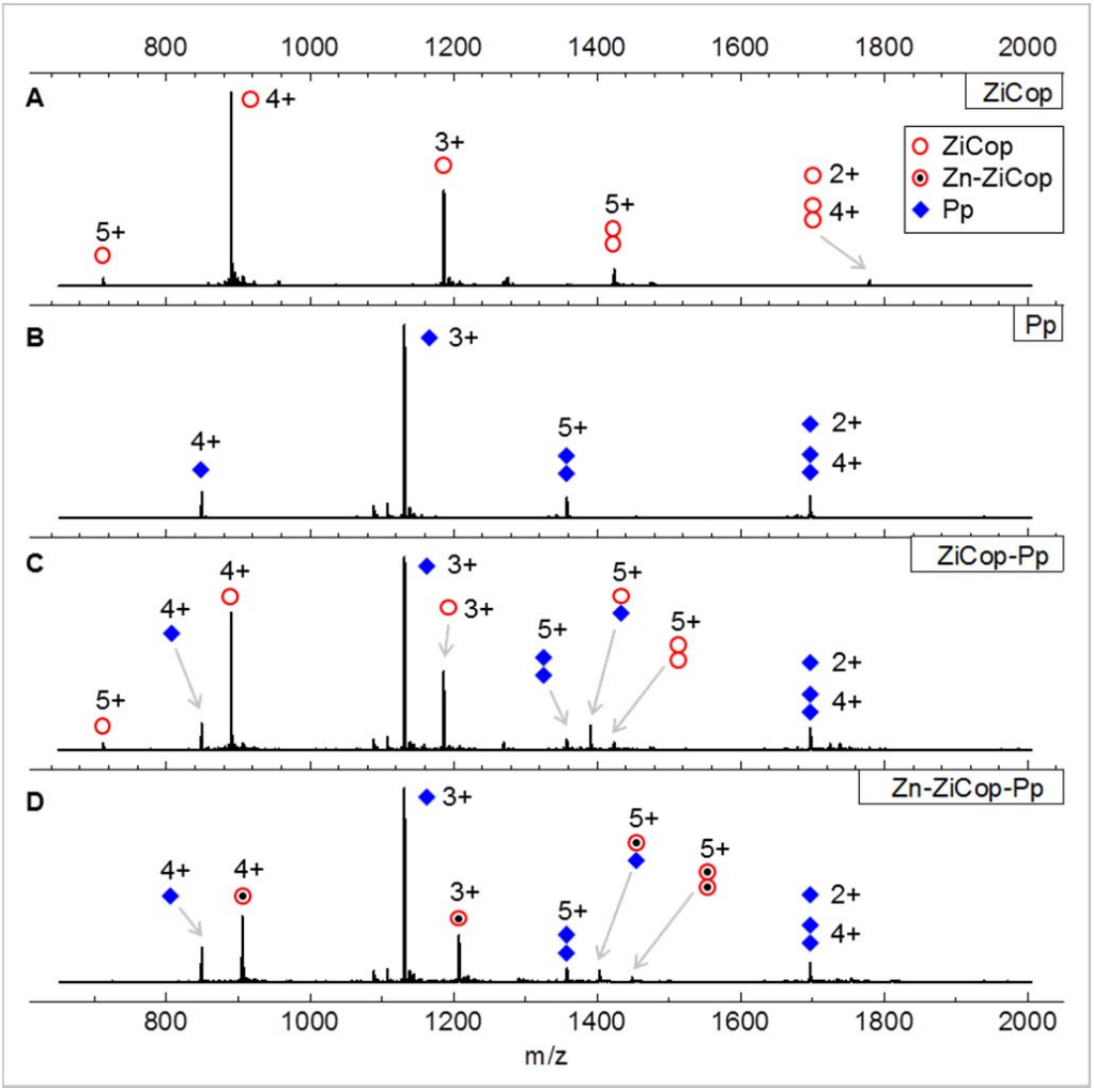


Figure 4-2 *Representative nESI mass spectra of ZiCop and partner peptide (Pp) obtained individually and mixed in the absence and presence of zinc. (**A**) and (**B**) – ZiCop and Pp respectively sprayed from the buffered conditions (see below); (**C**) – equimolar mixture of ZiCop and Pp, featuring monomeric and dimeric species; (**D**) – equimolar mixture of ZiCop, Pp and zinc acetate featuring full zinc binding by ZiCop and aggregation events. Conditions: 50 μM peptides and zinc acetate, 20 mM ammonium acetate, 5% isopropanol, 500 μM TCEP, pH 7.2. The peaks are denoted as follows (see legend in pane **A**: one for all spectra): apo-ZiCop by red circles, Zn-bound ZiCop by red circles with a black dot in the centre, Pp by blue diamonds, and any dimeric species are shown by clusters of relevant shapes.*

4.4 Metal ions and ZiCop – quantifying metal ion affinity

Figure 4-3 shows typical nESI mass spectra obtained for ZiCop without (***Figure 4-3A***) and with (***Figure 4-3B-D***) equimolar amounts of metal salts added; the regions for the 4+ charge states of the peptide are shown with an insert zoomed to show isotopic resolution. The $[ZiCop+4H]^{4+}$ or the equally charged $[ZiCop+X+2H]^{4+}$, where X = divalent cation, are the dominant peaks in the spectra, twice as intense as $[ZiCop+3H]^{3+}$. Use of TCEP (tris(2-carboxyethyl)-phosphine) maintained approximately 99% of the cysteines in their reduced state as evidenced by isotopic cluster analysis. Theoretical fits to the elemental compositions of the fully reduced peptide are superimposed on the experimental isotopic distribution (inserted zooms), which allows us to conclude that under these conditions the metals are coordinated by thiolate groups ($–S^-$).

The ratio between the intensity of the ^{12}C peak for both apo and holo species was determined for each spectrum, which provides an estimated order of metal affinity for apo-ZiCop, $Co^{2+} > Zn^{2+} >> Ca^{2+}$. Just under 20% of ZiCop is bound in the presence of Zn^{2+} (***Figure 4-3B***) and a quarter of the peptide population remains unbound in the presence of Co^{2+} (***Figure 4-3C***). A very low fraction of ZiCop (less than 10%) is associated with Ca^{2+} (***Figure 4-3D***). Addition of either Zn^{2+} or Co^{2+} narrows the charge state distribution of ZiCop in favour of the 4+ charge state (spectra not shown), which suggests metal-induced stabilisation of the fold.

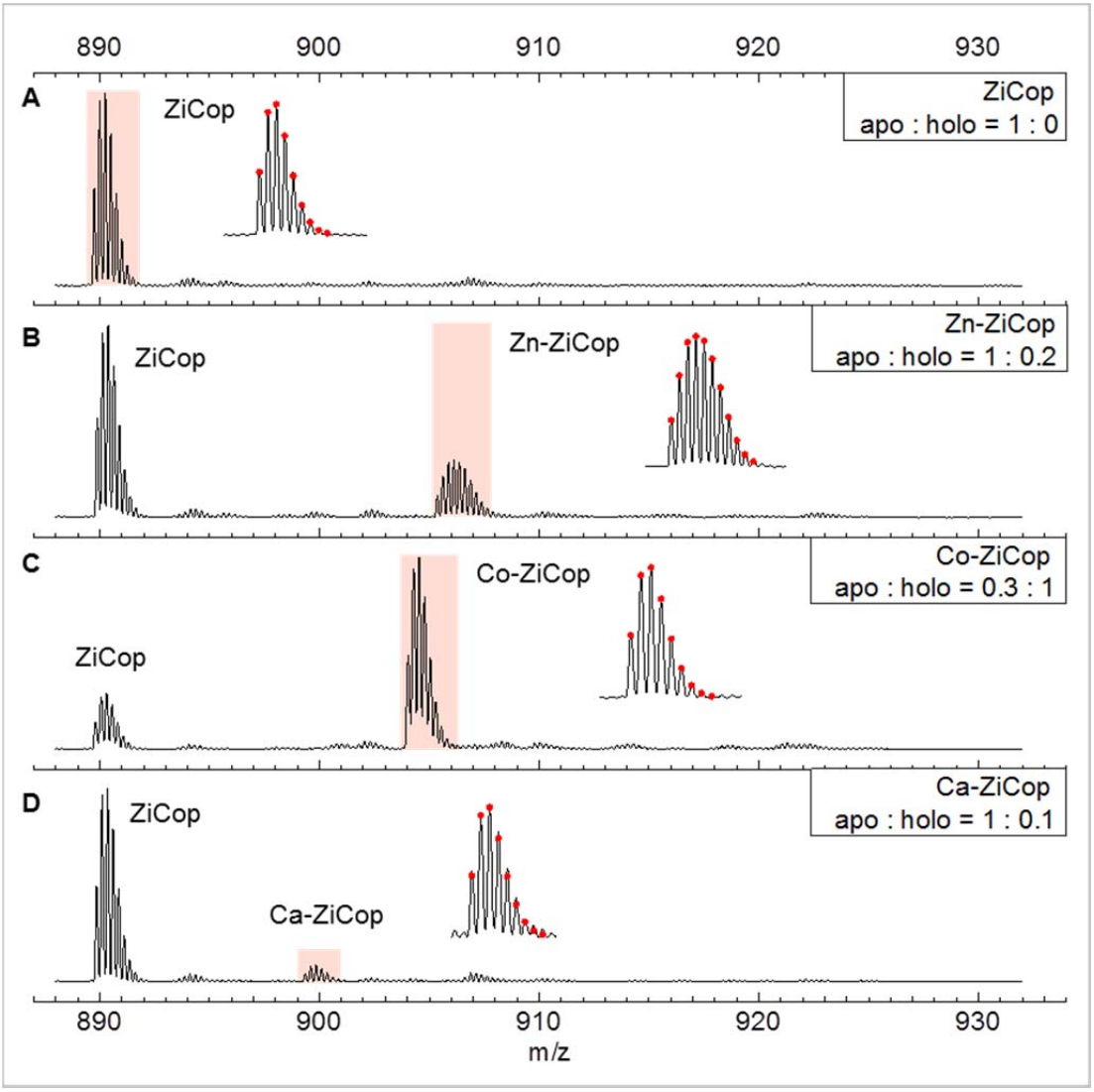


Figure 4-3 *Representative nESI mass spectra of the 4+ charge state of ZiCop for: (**A**) the apo state, and with: (**B**) Zn^{2+}, (**C**) Co^{2+} and (**D**) Ca^{2+}. Ratios of apo to holo form are calculated from the intensities of the mono-isotopic peaks. Insets show the resolved isotopic clusters for the free and bound states (highlighted in pink on the mass spectra) along with theoretical fitting (denoted by red dots). Isotopic cluster analysis indicated that 99% of the peptide is reduced. Conditions: 50 µM peptide; 20 mM ammonium acetate; 5% isopropanol; pH 7.2; 500 µM TCEP; 50 µM metal acetate salts.*

In order to determine K_d values for metal binding, all charge states were considered and the approach implemented for the vCP1 system in section ***3.4*** was applied here. ***Figure 4-4A-C*** shows binding curves obtained for ZiCop complexed by Zn, Co, and Ca respectively. In good agreement with the results shown in ***Figure 4-3***, strong binding is observed for Zn^{2+} and Co^{2+}

(with K_ds of 41 μM ± 2 and 17 ± 3 μM, respectively), and very poor for Ca^{2+} (K_d = 714 ± 45 μM). The sigmoidal shape of the zinc- and cobalt-binding curves, as well as the value of $n > 1$, suggests switching behaviour of ZiCop when presented with Zn^{2+} or Co^{2+} cation. ***Figure 4-4D*** shows the binding curve obtained by titrating Zn^{2+} into ZiCop equilibrated with an equimolar concentration of Co^{2+}. Competitive titrations between Zn^{2+} and Co^{2+} revealed a stronger zinc affinity to ZiCop (K_d = 2 ± 0.3 μM) compared to cobalt, than for Zn^{2+} titrated individually. This suggests not only that Zn^{2+} easily displaces Co^{2+}, but also that the presence of the latter 'configures' the binding site of the ZiCop for coordinating Zn^{2+}. Although the curve has a hyperbolic shape, the value of $n > 1$ indicates switching behaviour at the concentrations below the detection limit of the method employed. The competition experiment performed the 'opposite' way, *i.e.* by titrating cobalt into zinc-bound ZiCop in 1:1 stoichiometry reveals that no Zn^{2+} was displaced by Co^{2+} even at 5-fold molar excess of cobalt acetate relative to ZiCop and zinc acetate. This is strong evidence that ZiCop preferentially and specifically binds Zn^{2+} in the presence of Co^{2+}. Previously reported ESI-MS studies demonstrated that the order of affinity of different metals for calmodulin was altered in the presence of calcium[18].

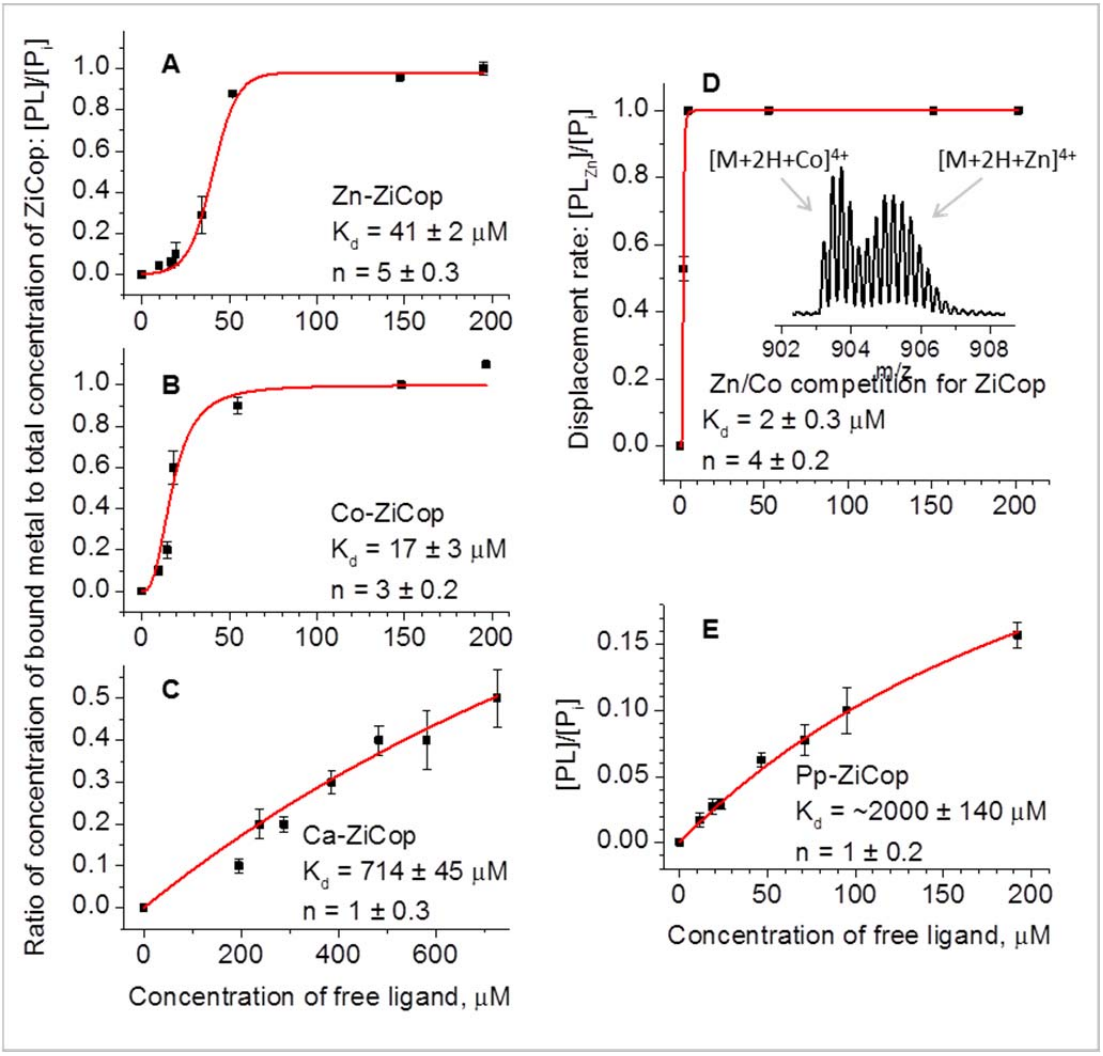


Figure 4-4 *Titration curves of ZiCop against metals and Partner Peptide obtained from titration nESI-MS experiments: (**A**) Zn^{2+}, (**B**) Co^{2+}, (**C**) Ca^{2+}, (**D**) displacement of Co^{2+} by Zn^{2+} from 50 μM ZiCop (inset - ZiCop : CoAc : ZnAc = 1 : 1 : 0.5, 4+ charge state), (**E**) Pp. In case of the metals and the Pp (**A** through **C** & **E**), the K_ds are determined by keeping the ZiCop concentration at 50 μM and titrating the metals and Pp in the following concentration ranges: Zn and Co from 0 to 250 μM; Ca from 0 to 750 μM; Pp from 0 to 200 μM. For the metal displacement experiments (**D**), the K_d was determined by keeping the ZiCop and CoAc concentrations at 50 μM and titrating ZnAc from 0 to 200 μM. Hill functions are fitted to the data points and the K_d values obtained at 50% system saturation, i.e. half of the peptide population binds metal. The K_d values and number of cooperative sites n is quoted on each panel. Data are derived from relative ion currents in ESI mass spectra (first monoisotopic peak), summing the intensities of ion currents for all charge states of each species. Each data point is the mean (± standard error) of the equilibrium concentrations from three mass spectra.*

Visible light absorbance measurements (carried out by the collaborator, see ***Appendix, Figure A 10B***) based on the displacement of cobalt by zinc, confirm

that ZiCop coordinates cobalt using a Cys_2His_2 coordination geometry, typical of classical zinc fingers[19; 20], with a K_d of around 7.5 μM. Zinc was found to displace cobalt from ZiCop, although the binding stoichiometry was found to be around 1.2 : 1 peptide:metal, suggesting that some of the peptide was oligomerising in the presence of metal. ZiCop was found to bind cobalt with lower affinity in the presence of partner peptide, but it is not known if this is due to competition between the coiled coil and zinc-bound states, and how much is due to the conformation adopted when partner peptide is interacting with the zinc-bound ZiCop. The analysis of shape of the absorbance spectrum (***Appendix, Figure A 10A***) in the presence of partner peptide indicates that the mode of cobalt binding was unchanged.

4.5 ZiCop and Partner Peptide – quantifying the strength of interaction

For ZiCop interacting with Pp, the first assumption made above (section ***3.4***) regarding the K_d measurements in the gas phase is not entirely adhered to, as the masses of both interacting partners are now very similar. However, this is well compensated by the fact that ionisation efficiencies of both (as monomers) are very similar too, so one can surmise contribution of each of them is equal. This has been confirmed by the summation of peak intensities across all charge states for each of the peptide on the mass spectrum that featured a mixture of equimolar amounts of ZiCop and Pp sprayed into the mass spectrometer. ***Figure 4-4E*** shows the binding curve of ZiCop with Pp which is at least two orders of magnitude weaker than that found for Zn^{2+} and Co^{2+}. This is not unexpected given that coiled coil associations are

innately weaker than those of coordinated metals[21; 22]. Extrapolation of the Hill curve gives the value of $K_d \approx 2 \pm 0.1$ mM. The partner peptide was also found to interact with the Zn-bound form of ZiCop, a feature that was also observed by circular dichroism (data collected and analysed by the collaborator; not shown).

4.6 Peptide-metal and peptide-peptide – qualitative definition of complex stability by CID

Collision induced dissociation (CID) can be employed to probe the stability of non-covalent complexes and give an indication of specificity of non-covalent interactions[23; 24]. Considered here are the following 5+ charge state complex ions: Zn-bound form of ZiCop interacting with Pp – Zn-ZiCop-Pp (*m*/*z* 1403.3); hetero-dimer of the ZiCop and Pp – ZiCop-Pp (*m*/*z* 1391.0); two homo-dimers of both peptides – 2ZiCop (*m*/*z* 1423.9) and 2Pp (*m*/*z* 1358.1); dimer of the Zn-bound form of ZiCop – 2(Zn-ZiCop) (*m*/*z* 1448.5). The Zn-ZiCop-Pp and 2(Zn-ZiCop) are the 'artefacts' of the design, which were not anticipated but are observed, and therefore are of particular interest in terms of understanding the system and improving future designs. The ZiCop-Pp was designed to form a hetero-dimeric coiled coil ahead of the two homo-dimeric colied coils, hence anticipated differences in relative dissociation energies will confirm the success of the design strategy (***Figure 4-1***). As the relative dissociation energy is increased in 5 V increments (***Figure 4-5***), the complexes (green curves) dissociate into their constituent parts yielding

unique values of voltage at 50% precursor population dissociation (E_{50})[25; 26]. These values are used to gauge the strength of complex association: the more energy is required to break up the interaction the more stable the complex is.

The strongest associations are observed for ZiCop interacting with its partner – in both Zn-bound and free forms (highest E_{50} values of 21 and 16 V respectively, ***Figure 4-5A*** and ***B***). The stronger association between the two peptides in the presence of zinc ion is presumably a consequence of the metal-stabilised fold of the ZiCop. For the Zn-ZiCop-Pp complex, the two major products are triply-charged Zn-ZiCop and doubly-charged Pp, in case of ZiCop-Pp complex – triply-charged ZiCop and doubly-charged Pp. Additional experiments (data not shown) revealed that loss of Zn^{2+} was an unobservable channel, the peptide fragmented before the loss of the metal. This confirms the specific and strong interaction of zinc in ZiCop when in the Zn-ZiCop-Pp complex, and it can be concluded that the interaction between Zn-ZiCop and Pp occurs *via* the α-helical part of the zinc finger fold of the ZiCop.

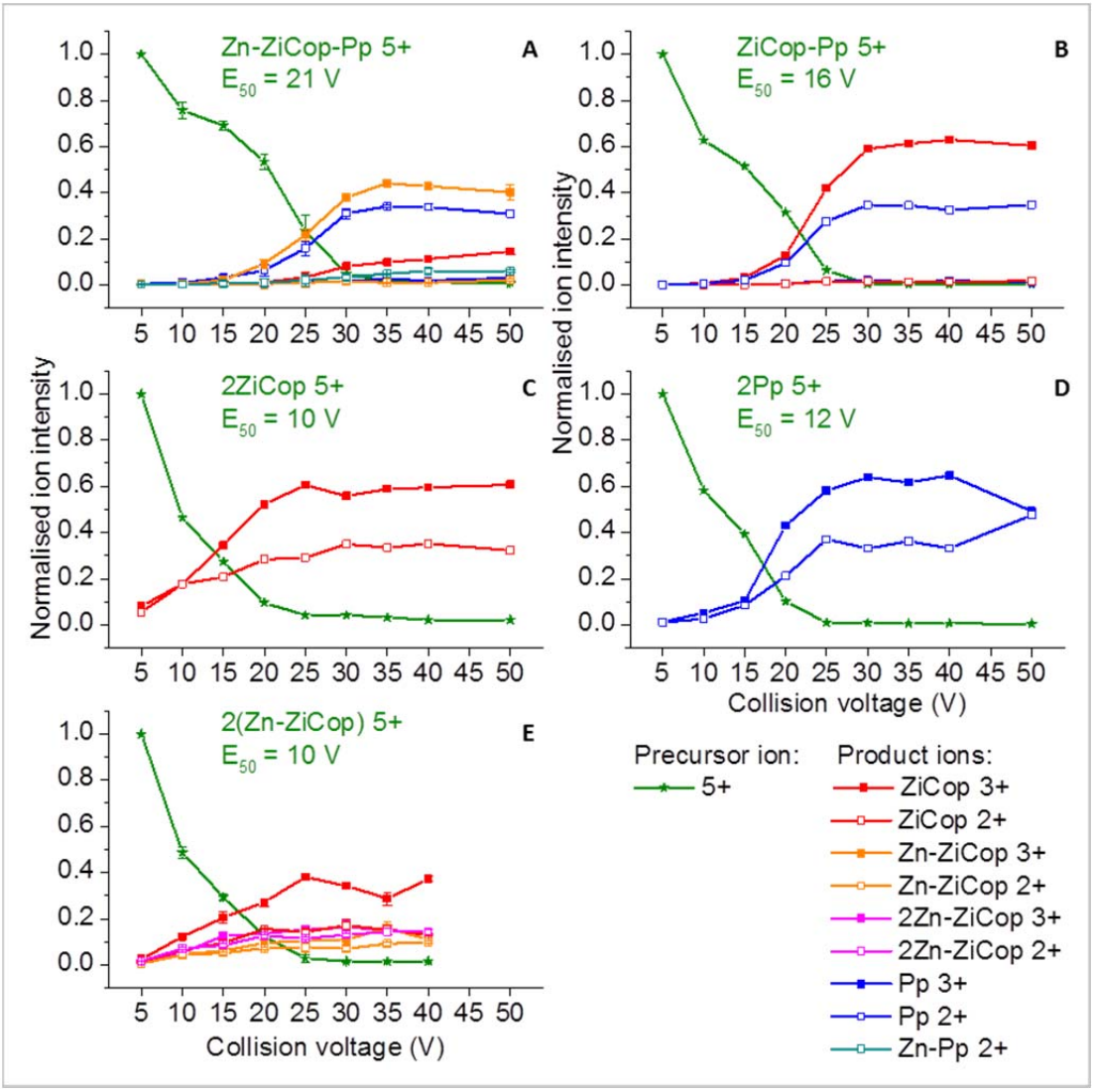


Figure 4-5 *Stability of non-covalent complexes as a function of dissociation energy. Fragmentation results of the following 5+ charge state complexes are shown: (**A**) – Zn-ZiCop-Pp; (**B**) – ZiCop-Pp; (**C**) – 2ZiCop; (**D**) – 2Pp; (**E**) – 2(Zn-ZiCop). Stability of complexes is calculated from the intensities of the precursor and product ions during CID. The green curves show dissociation of the precursor ion, all other curves are dissociation products, and the E_{50} values for the precursors are shown for each plot. Each data point is the mean (± standard error) of the equilibrium concentrations from three mass spectra.*

Comparative analysis of the dissociation curves of the metal-free hetero-dimer (***Figure 4-5B***) and two homo-dimers (***Figure 4-5C*** and ***D***) reveal that the strongest association is observed for the ZiCop-Pp ion (E_{50} = 16 V), followed by the 2Pp (E_{50} = 12 V) and 2ZiCop (E_{50} = 10 V) homo-dimers. The same order affinity was observed by spectroscopic methods (not shown). The

dissociation product of the dimerised zinc-bound ZiCop – 2(Zn-ZiCop) shown in ***Figure 4-5E*** – is largely peptide apo-form, with a small population of holo-ZiCop. Thus one can infer that a large proportion of the precursor ion was not involved in specific association between peptide and metal. Additionally, the 2(Zn-ZiCop) ion has the same dissociation energy as the 2ZiCop ion, yet again confirming the non-specific nature of the 2(Zn-ZiCop) aggregate whereby the metal ion is not coordinated tetrahedrally. Indeed, stabilisation of coiled coils by metal ions has been explored as an independent target design by a number of groups, whereby Cys and/or His on the first coil and the same residues on the second (and sometimes third) coil in the bundle are held together by transition metal ions[27; 28]. The K_ds of such associations are one to two orders of magnitude higher[28; 29] (*i.e.* weaker affinity) than those reported for designed zinc finger folds measured by analogous spectroscopic methods[15; 30]. Therefore, such metal-sandwiched complexes are innately weaker than highly-structured zinc finger folds, and the observed effect here is possibly due to this type of associations.

These findings confirm that the adopted peptide design strategy yielded the desired preference for the hetero-dimer formation over either of the homo-dimers. Moreover, the mass-spectrometry based platform is selective enough to distinguish the interaction energies between the ZiCop and Pp self-association, with the latter having no sequence duality and hence a slightly stronger propensity for coiled coil formation compared to the former.

4.7 Collision cross sections and molecular dynamics simulations – elucidating conformations

IM-MS experiments were performed on the apo and the holo forms of ZiCop and its complexes with Pp. Arrival time distributions (ATDs) were obtained at a range of drift voltages which were converted to collision cross sections following ***Equation 2-5*** in ***Chapter 2***, and the CCS values are presented in ***Table 4-1***. For ZiCop, the cross section difference between the two charge states was around 20%, in agreement with the ion cross section dependency on the number of charges it carries: Coulombic effects due to proximal charges are well known to induce unfolding in gas phase peptides and proteins[31; 32; 33]. By contrast, the CCS for the two charge states of zinc bound ZiCop were in good agreement (***Table 4-1A***). This can be attributed to structural stability rendered by the metal ion. The CCS found for the hetero-dimer suggests an elongated structure that is significantly larger than either of the two monomer species (***Table 4-1B***). This is also the case for both homo-dimers, which had very similar cross sections. The dimer of the ZiCop, however, gave a larger than expected standard deviation of experimental Ω values, suggesting conformational variance, possibly due to multiple inter-converting populations. This may be related to the fact that ZiCop is a very poor coiled coil, which was the initial reason for introducing a binding partner – Pp. The 5+ Zn-ZiCop-Pp ion has a larger Ω value than any of the non-metal dimers, which supports this complex as having an interface between the α-helical part of the zinc finger and the helix of the partner peptide. This is corroborated by the CID results. The 5+ dimer of Zn-ZiCop has a very similar collision cross section to the Zn-ZiCop-Pp, however, there

is no evidence by CID that the ZiCop forms a zinc finger fold within this arrangement. The most plausible spatial arrangement of the 2(Zn-ZiCop) is the one that involves globular conformations of the two ZiCop chains interacting with each other *via* hydrophobic residues, and neither of the two zinc ions coordinating the cysteines and histidines, but rather interacting with the charged side chains electrostatically.

Charge	ZiCop		Zn-ZiCop	
	IM-MS	MD	IM-MS	MD
3+	558 ± 1	*550*	648 ± 3	*645*
4+	664 ± 6	*641*	656 ± 4	

A

Charge	ZiCop-Pp		2Pp	2ZiCop	Zn-ZiCop-Pp	2(Zn-ZiCop)
	IM-MS	MD	IM-MS	IM-MS	IM-MS	IM-MS
5+	971 ± 24		975 ± 3	947 ± 118	1104 ± 61	1103 ± 13
6+		*1007*				

B

Table 4-1 *Experimental (IM-MS) and simulated (MD) collision cross sections of the most abundant charge states of the monomeric apo and holo-ZiCop (Zn-bound) – table (**A**), the dimeric species, and some of the undesirable aggregates – table (**B**), in 'buffered' conditions. The values are quoted in Å^2, with standard error of three experimental repeats. The values for the modelled structures are shown in italics. The empty cells denote values that were not determined.*

To investigate these phenomena further, molecular dynamics (MD) simulations were performed on selected species shown in ***Table 4-1***: apo- and holo-ZiCop and the hetero-dimer of the two – ZiCop-Pp. ***Figure 4-6*** shows their representative structures obtained by MD simulations. Experimental collision cross sections compared favourably with those obtained from the MD giving confidence in both datasets. Interestingly, the apo-ZiCop adopted a very similar conformation to that of the holo-form (***Figure 4-6A*** and ***B***), with the main difference observed in the coordination sphere: the Cys and His residue side chains seem to be on the outer surface of the peptide globule in the absence of the metal ion. The α-helical portion

of the peptide chain tends to be present in both apo- and holo-forms. The hetero-dimer (***Figure 4-6C***) adopts an elongated conformation consisting of the two parallel α-helices slightly twisting around each other. Although in the resulting coiled coil the termini of both chains are somewhat unstructured, the alignment of the peptides is blunt-ended as envisioned by the design.

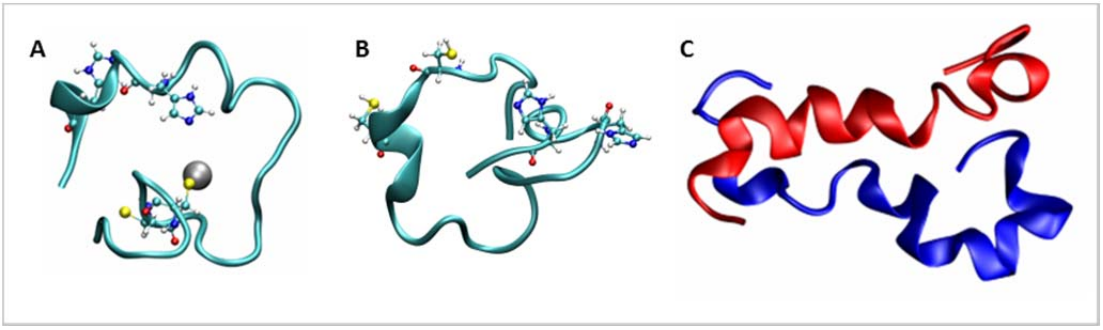


Figure 4-6 *Representative MD structures of the 3+ Zn-ZiCop (**A**), 3+ apo-ZiCop (**B**) and 6+ ZiCop-Pp hetero-dimer (**C**). Metal-coordinating residues are shown as ball-and-stick representations to highlight their role in the presence (**A**) and absence (**B**) of zinc ion. Hetero-dimer (**C**) is colour-coded according to the convention in the present work to distinguish between the two interacting partners.*

4.8 Conclusions

A mass spectrometry and ion-mobility mass spectrometry workflow to interrogate a protein design strategy has been presented. The following are this work's primary findings:

❖ There is strong evidence that ZiCop preferentially and specifically binds Co^{2+} and, to a lesser extent Zn^{2+}, when compared to and Ca^{2+}. The presence of Co^{2+}, before Zn^{2+} is added to the peptide, enhances the strength of

its binding by at least an order of magnitude. The shape of binding curve of Co^{2+} is sigmoidal and is similar to that of Zn^{2+}, suggesting switching behaviour of the peptide.

❖ Collision cross sections (CCS) measurements revealed an increased conformational stability of the zinc-bound ZiCop over its apo-form.

❖ Binding of Ca^{2+} is very weak and non-specific, with the binding curve taking a hyperbolic shape, indicating non-switching behaviour.

❖ Dissociation constant of hetero-dimer ZiCop-Pp formation is in the low-mM range. IM-MS measurements and MD simulations indicate the possibility that ZiCop-Pp species adopt a blunt-ended coiled coil conformation. The hetero-dimer forms preferentially over the two homo-dimers – 2ZiCop and 2Pp.

4.9 References

1. Socolich, M., Lockless, S. W., Russ, W. P., Lee, H., Gardner, K. H. & Ranganathan, R. (2005). Evolutionary information for specifying a protein fold. *Nature (London, U. K.)* **437**, 512-518.

2. Jin, W., Kambara, O., Sasakawa, H., Tamura, A. & Takada, S. (2003). De Novo Design of Foldable Proteins with Smooth Folding Funnel Automated Negative Design and Experimental Verification. *Structure (Cambridge, MA, U. S.)* **11**, 581-590.

3. Tuinstra, R. L., Peterson, F. C., Kutlesa, S., Elgin, E. S., Kron, M. A. & Volkman, B. F. (2008). Interconversion between two unrelated protein folds in the lymphotactin native state. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5057-5062.

4. Yan, H. & Tsai, M.-D. (1999). Nucleoside monophosphate kinases: structure, mechanism, and substrate specificity. *Adv. Enzymol. Relat. Areas Mol. Biol.* **73**, x-xi, 103-134, 4 plates.

5. Ye, Q., Wang, H., Zheng, J., Wei, Q. & Jia, Z. (2008). The complex structure of calmodulin bound to a calcineurin peptide. *Proteins: Struct., Funct., Bioinf.* **73**, 19-27.

6. Kelly, J. W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* **8**, 101-106.

7. Berezovskaya, Y., Armstrong, C. T., Boyle, A. L., Porrini, M., Woolfson, D. N. & Barran, P. E. (2011). Metal binding to a zinc-finger peptide: a comparison between solution and the gas phase. *Chem. Commun.* **47**, 412-414.

8. Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. (1995). Naked Protein Conformations - Cytochrome-C in the Gas-Phase. *J. Am. Chem. Soc.* **117**, 10141-10142.

9. Wyttenbach, T. & Bowers, M. T. (2007). Intermolecular Interactions in Biomolecular Systems Examined by Mass Spectrometry. *Annu. Rev. Phys. Chem.* **58**, 511-533.

10. Mesleh, M. F., Hunter, J. M., Shvartsburg, A. A., Schatz, G. C. & Jarrold, M. F. (1996). Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. *J. Phys. Chem.* **100**, 16082-16086.

11. Shvartsburg, A. A. & Jarrold, M. F. (1996). An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86-91.

12. Wyttenbach, T., vonHelden, G., Batka, J. J., Carlat, D. & Bowers, M. T. (1997). Effect of the long-range potential on ion mobility measurements. *J. Am. Soc. Mass. Spectrom.* **8**, 275-282.

13. Valentine, S. J., Counterman, A. E. & Clemmer, D. E. (1999). A database of 660 peptide ion cross sections: use of intrinsic size parameters for bona fide predictions of cross sections. *J. Am. Soc. Mass. Spectrom.* **10**, 1188-1211.

14. De Cecco, M., Seo, E. S., Clarke, D. J., McCullough, B. J., Taylor, K., Macmillan, D., Dorin, J. R., Campopiano, D. J. & Barran, P. E. (2010). Conformational Preferences of Linear OI-Defensins Are Revealed by Ion Mobility-Mass Spectrometry. *J. Phys. Chem. B* **114**, 2312-2318.

15. Cerasoli, E., Sharpe, B. K. & Woolfson, D. N. (2005). ZiCo: A peptide designed to switch folded state upon binding zinc. *J. Am. Chem. Soc.* **127**, 15008-15009.

16. Acharya, A., Rishi, V. & Vinson, C. (2006). Stability of 100 Homo and Heterotypic Coiled-Coil a-a' Pairs for Ten Amino Acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry* **45**, 11324-11332.

17. Krylov, D., Barchi, J. & Vinson, C. (1998). Inter-helical interactions in the leucine zipper coiled coil dimer: ph and salt dependence of coupling energy between charged amino acids. *J. Mol. Biol.* **279**, 959-972.

18. Shirran, S. L. & Barran, P. E. (2009). The Use of ESI-MS to Probe the Binding of Divalent Cations to Calmodulin. *J. Am. Soc. Mass. Spectrom.* **20**, 1159-1171.

19. Laity, J. H., Lee, B. M. & Wright, P. E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **11**, 39-46.

20. Krishna, S. S., Majumdar, I. & Grishin, N. V. (2003). Structural classification of zinc fingers: survey and summary. *Nucl. Acids Res.* **31**, 532-550.

21. Litowski, J. R. & Hodges, R. S. (2002). Designing Heterodimeric Two-stranded α-Helical Coiled-coils. *J. Biol. Chem.* **277**, 37272-37279.

22. Reddi, A. R., Guzman, T. R., Breece, R. M., Tierney, D. L. & Gibney, B. R. (2007). Deducing the Energetic Cost of Protein Folding in Zinc Finger Proteins Using Designed Metallopeptides. *J. Am. Chem. Soc.* **129**, 12815-12827.

23. Smith, R. D., Bruce, J. E., Wu, Q. & Lei, Q. P. (1997). New mass spectrometric methods for the study of noncovalent associations of biopolymers. *Chem. Soc. Rev.* **26**, 191-202.

24. Heck, A. J. R. & Jørgensen, T. J. D. (2004). Vancomycin in vacuo. *Int. J. Mass Spectrom.* **236**, 11-23.

25. Benesch, J. L. P., Aquilina, J. A., Ruotolo, B. T., Sobott, F. & Robinson, C. V. (2006). Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chem. Biol.* **13**, 597-605.

26. Robinson, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M. & Dobson, C. M. (1996). Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein - CoA Ligand Binding and Assembly. *J. Am. Chem. Soc.* **118**, 8646-8653.

27. Tanaka, T., Mizuno, T., Fukui, S., Hiroaki, H., Oku, J., Kanaori, K., Tajima, K. & Shirakawa, M. (2004). Two-Metal Ion, Ni(II) and Cu(II),

Binding α-Helical Coiled Coil Peptide. *J. Am. Chem. Soc.* **126**, 14023-14028.

28. Ghosh, D. & Pecoraro, V. L. (2005). Probing metal-protein interactions using a de novo design approach. *Curr. Opin. Chem. Biol.* **9**, 97-103.

29. Ghosh, D. & Pecoraro, V. L. (2004). Understanding Metalloprotein Folding Using a de Novo Design Strategy. *Inorg. Chem.* **43**, 7902-7915.

30. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L. & Berg, J. M. (1991). A consensus zinc finger peptide: design, high-affinity metal binding, a pH-dependent structure, and a His to Cys sequence variant. *J. Am. Chem. Soc.* **113**, 4518-4523.

31. Shelimov, K. B., Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. (1997). Protein Structure in Vacuo: Gas-Phase Conformations of BPTI and Cytochrome c. *J. Am. Chem. Soc.* **119**, 2240-2248.

32. Valentine, S. J., Anderson, J. G., Ellington, A. D. & Clemmer, D. E. (1997). Disulfide-Intact and -Reduced Lysozyme in the Gas Phase: Conformations and Pathways of Folding and Unfolding. *J. Phys. Chem. B* **101**, 3891-3900.

33. Covey, T. & Douglas, D. J. (1993). Collision cross sections for protein ions. *J. Am. Soc. Mass Spectrom.* **4**, 616-23.

Chapter 5

The effect of salt on protein conformation and stability

5.1 Introduction

This chapter describes the effect of salt adductation on the conformations of three model proteins of varying molecular mass (lysozyme, cytochrome c and BPTI) by variable temperature ion mobility mass spectrometry. The proteins are incubated with sodium iodide and following nESI, their gas-phase conformations are determined using drift tube ion mobility mass spectrometry (DT IM-MS) at three different drift cell temperatures – 'ambient' (300 K), 'cold' (260 K) and 'hot' (360 K). Significant adductation of hydrogen iodide is observed on all three proteins with the number of HI molecules correlating to the number of available basic sites. The conformational space occupied by each protein is reduced significantly in the presence of salt. Thermally induced unfolding (which is observed at both cold and hot temperatures) is minimised in the presence of adducted HI. This is the first time the phenomenon of 'cold denaturation' is observed for proteins in the gas phase, suggesting that this effect is intrinsic to the protein fold.

Protein-salt interactions play a key role in the structure of a protein and on stabilising its active form. Franz Hofmeister was the first to consider how salt ions affect proteins, and this work culminated in the 1888 milestone publication 'Zur Lehre von der Wirkung der Salze' (About the Science of the Effect of Salts)[1]. He found that these effects vary widely, however some trends could be established for cations and anions separately. These trends were the salts' ability to precipitate proteins, which gave rise to the Hofmeister series – which contained a separate ranking for cations and anions. Recent years have seen a revival of curiosity in this topic with a more than 3-fold increase in citations in the past decade compared to the previous one. This increased interest in the Hofmeister effect is testament to its fundamental relevance to a wide range of fields: from enzyme activity[2] and protein stability[3] to protein-protein interactions[4] and protein crystallisation[5] and also to the development of new experimental tools[6] and models[7] that can help shed light on this long-established phenomenon. As a result of these recent research efforts, the initial hypothesis that the Hofmeister phenomenon is due to the effect of salts in 'making' or 'breaking' the structure of water has gradually fallen out of favour[8]. In its place has come an assertion that the key instigator of precipitation is *via* direct ion-macromolecule interactions. These interactions are very amenable to gas-phase studies.

Through introduction of soft ionisation techniques of MALDI[9; 10] and ESI[11; 12], mass spectrometry has become a popular technology underpinning protein research. Gas-phase protein ions[13] and their complexes[14; 15; 16] are now studied directly as a population of 'naked' ions[17], decoupled from any counter-ions. Volatile buffers are used to provide 'clean' mass spectra[18] although native-ESI mass spectrometry peaks are often somewhat broadened by the presence

of un-desolvated buffer salts and/or water[19]. Further structural elucidation is enabled by collision induced dissociation (CID)[20], whereby protein complexes are dissected into their constituent parts[21; 22; 23] by collisions with inert gas. However, in biological systems, proteins are never alone: they are crowded by other macromolecules and ions which play a crucial role in their stabilisation and function[24]. The focus of this work is to examine the effect of anions on proteins with mass spectrometry. In recent years, mass spectrometry based studies, and in particular those which aim to preserve biologically relevant conformations for gas-phase analysis, have gradually evolved from considering adducted salts as a regrettable feature to regarding them key to retaining a stable gas-phase fold [25; 26]. It has been shown recently that anions attached to proteins tend to stabilise compact conformations[26] and that the number of certain anions (such as perchlorate and iodide) that can adduct to a protein, correlates well with the number of available basic sites[27]. It follows from this that a protein surface can be 'mapped' by such anions.

All three proteins investigated here are well characterised by other biophysical methods and therefore are good models for this study. Their crystal structures were first published in the 1970s and 1980s, and to date, more than 560 structures are available for lysozyme, over 200 for cytochrome *c* and approaching 100 for BPTI[28].

Measurable parameters in the gas phase to address this question are:

- Relative intensities of the signals due to bare protein ions compared with those from protein-salt complexes;

❖ Changes in collision cross sections of these ions as a function of the number of hydrogen iodide molecules attached to the protein (up to 3 are considered);

❖ Changes in collision cross sections of these ions as a function of the buffer gas temperature.

Ion mobility measurements were carried out on these proteins at three different drift cell temperatures – 'ambient' (300 K), 'cold' (260 K) and 'hot' (360 K) to probe the conformational dynamism of the proteins. They were sprayed from their aqueous solutions with a 40-fold molar excess of NaI, and their apo-form and hydrogen iodide adducts (up to 3 HI molecules) were investigated. In the presented work, the term 'adduct' is used in relation to hydrogen iodide, rather than sodium or sodium iodide. Theoretical CCS were calculated in MobCal for the crystal structure coordinates of the three proteins at these temperatures and subsequently compared to experimental values.

5.1.1 Choice of salt

Sodium iodide was chosen for this study for a number of reasons. Iodide is a chaotropic anion in the Hofmeister series[1] and, along with other ingredients, it is one of the pharmaceutically acceptable anions (*i.e.* used in formulations of medicines) used to stabilise biomolecules (proteins, DNA, RNA) in biological matrices (blood, saliva, urine *etc*.) for their prolonged storage and shipment[29]. Iodide has a destabilising effect on protein complexes[26], as it can bind to non-polar patches on proteins, in addition to cationic residues[30]. Of

all halides, iodide has the largest ionic radius (disregarding astatine, which is unstable and radioactive, and hence pharmaceutically unsuitable)[31], measuring to ~2.4 Å[32; 33] and therefore the lowest charge density. For this reason, this anion will bind to proteins more tightly[5] and any anticipated resulting conformational changes will be more profound and easier to observe. Finally, anions generally have a larger effect on proteins than cations[8; 25; 34] .

5.1.2 Choice of electrospray solvent conditions

A comparison of two solvent conditions has been carried out: aqueous and buffered (20 mM NH_4OAc). ***Figure 5-1*** shows the results of the spray solvent comparison for lysozyme. In order to stabilise proteins for ESI-MS, it is common to employ a volatile buffer such as ammonium acetate[35]. To assess the effect of NaI on proteins properly, lysozyme was examined from aqueous solution as well as from buffered solution. Here, 50 μM lysozyme was sprayed from pure water (***Figure 5-1A*** and ***B***) and buffer (***Figure 5-1C*** and ***D***), with (panes ***B*** and ***D***) and without (panes ***A*** and ***C***) addition of 40-fold molar excess of NaI. Two general features that are associated with buffering of the spray solution are observed here. Firstly, the charge state distribution is shifted towards lower values (dominant charge state 8+) compared to aqueous conditions (dominant charge state 9+). Secondly, in the absence of salt, buffering tightens the charge state envelope (*cf*. ***A*** and ***C***). Remarkably, addition of salt to the buffered solution (*cf*. ***C*** and ***D***) does not significantly change the shape of the charge state envelope, whereas in aqueous solution salt causes tightening of this distribution (*cf*. ***A*** and ***B***) bringing it closer to that of the corresponding buffered solution (*cf*. ***B*** and ***D***). Thus, NaI has a

similar effect on the compaction of protein charge state distribution to that of NH_4OAc, suggesting that the high ionic strength also reduces the conformational spread of the protein in solution, as a buffer would. At the same time, in the absence of NH_4OAc, fewer Na^+ cations are clustered on each of the I^--adducted states of lysozyme, decoupling the anion from the cation effect (*cf*. the zoomed-in regions in panels ***B*** and ***D***). Therefore, aqueous conditions are the spray solution of choice for the present investigation.

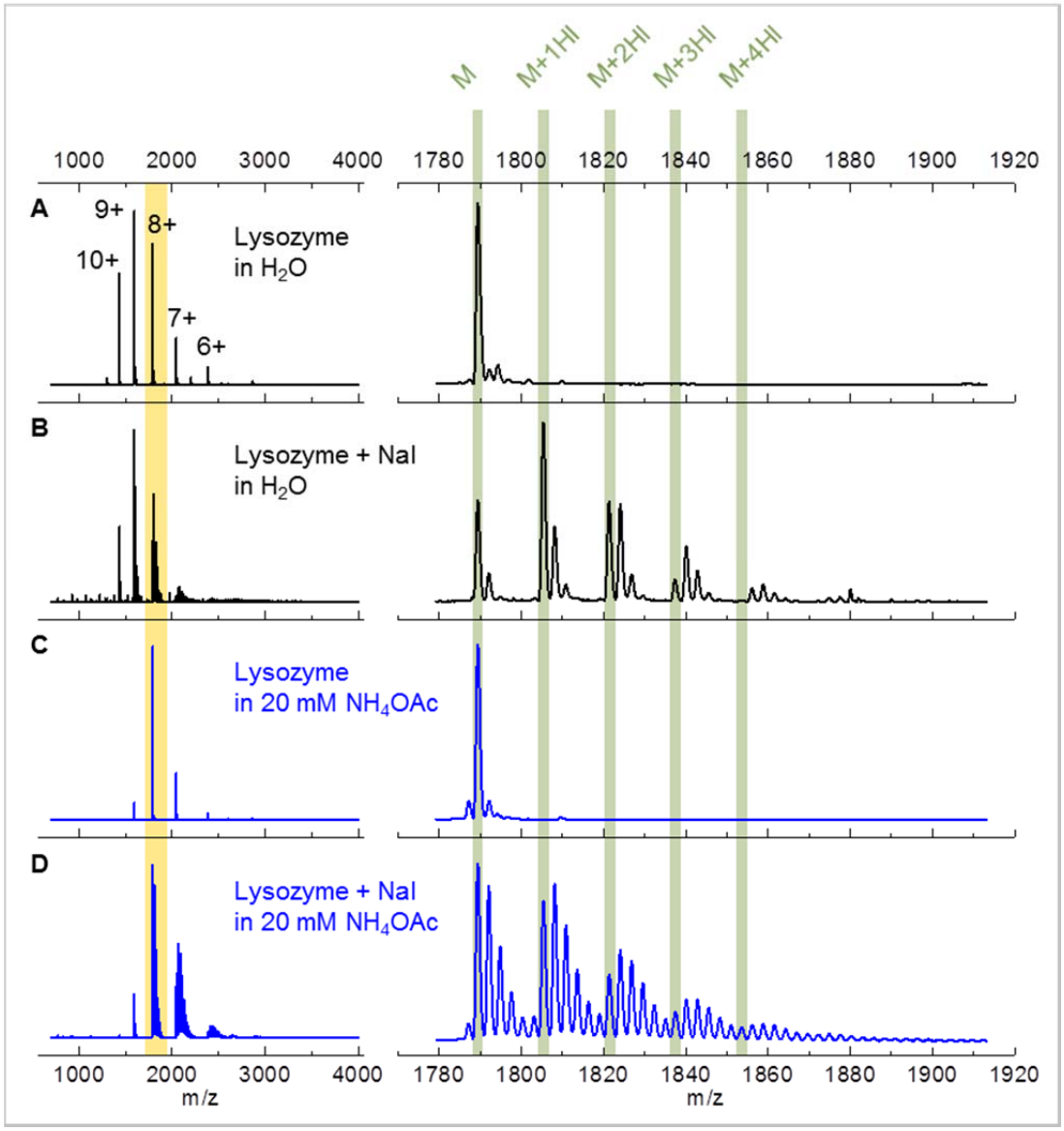


Figure 5-1 *Mass spectra of 50 μM lysozyme in aqueous, pH7.0 (panes **A** and **B**) and buffered, pH7.3 (20 mM NH_4OAc, panes **C** and **D**) conditions, with addition of 40-fold molar excess of NaI (panes **B** and **D**) and without salt (panes **A** and **C**). The data were acquired on the Q-ToF-2 instrument. Left column shows full-range mass spectra with the 8+ charge state highlighted in yellow; the right column illustrates the detailed view of the highlighted range of the 8+ peak. The 8+ ion of lysozyme and its hydrogen iodide adducts are highlighted in green. The peaks that follow the highlighted series are sodium clusters (panes **B** and **D**).*

5.1.3 Overview of sodium iodide interaction with proteins

Figure 5-2 shows representative mass spectra of the three proteins (lysozyme, cytochrome *c* and BPTI) sprayed from aqueous solution with a 40-fold molar excess of NaI. Lysozyme (***Figure 5-2A***) presents itself in 5 charge states – from 7+ to 11+, with the 10+ being most dominant, closely followed by the 9+. The first peak in each of the charge states is the MS signal of the apo-from of the protein, and the rest of the peaks are due to HI adduction: each next peak in the series is the previous with one additional hydrogen iodide (see inset in ***Figure 5-2A***). When hydrogen iodide adducts, the mass added is that of H + I, (127.9 Da.) with respect to the apo species although this method cannot distinguish where the anion has attached. The sodium adduct peaks correspond to an addition of 22.0 Da suggesting that they have replaced a proton. HI is also more dominant as an adduct than sodium (see insets in ***Figure 5-2***), although the ratio between bound NaI and bound HI increases as the number of iodides increases. In the case of NaI addition, the mass increase corresponds to 149.9 Da. For all three proteins the average number of retained salt ions increases as the charge state decreases, with a markedly decreased abundance of the apo species, suggesting that a fully salted out form of the protein would be electrostatically neutral, as in solution.

Cytochrome *c* (***Figure 5-2B***) has a very tight charge state distribution – from 6+ to 9+, with the two flanking charge states of very low abundance, and the two central ones of comparable intensity. Similarly to lysozyme, each charge state of cytochrome *c* presents itself as a cluster of hydrogen iodide adducts, with higher levels of adduction observed for lower charge states. Cytochrome *c* displays an especially high propensity to iodide clustering, as evidenced by the very low abundance of the apo-form and high number of

HI-adducted peaks (***Figure 5-2B*** inset). Here, the 7+ charge state features 8 HI adducts, with the 4-adducted peak being most populated. Clusters of peaks within each iodide adduction state are due to sodium ions.

BPTI, similarly to cytochrome *c*, has two almost equally dominant charge states (5+ and 6+) flanked by the low-abundance 4+ and 7+ peaks (***Figure 5-2C***). The apo-form is dominant for all charge states except for the 4+, and adduction level is very low for this protein compared to the other two (see inset in ***Figure 5-2C***). The trend that lower charge states retain more adducts holds true for this protein too.

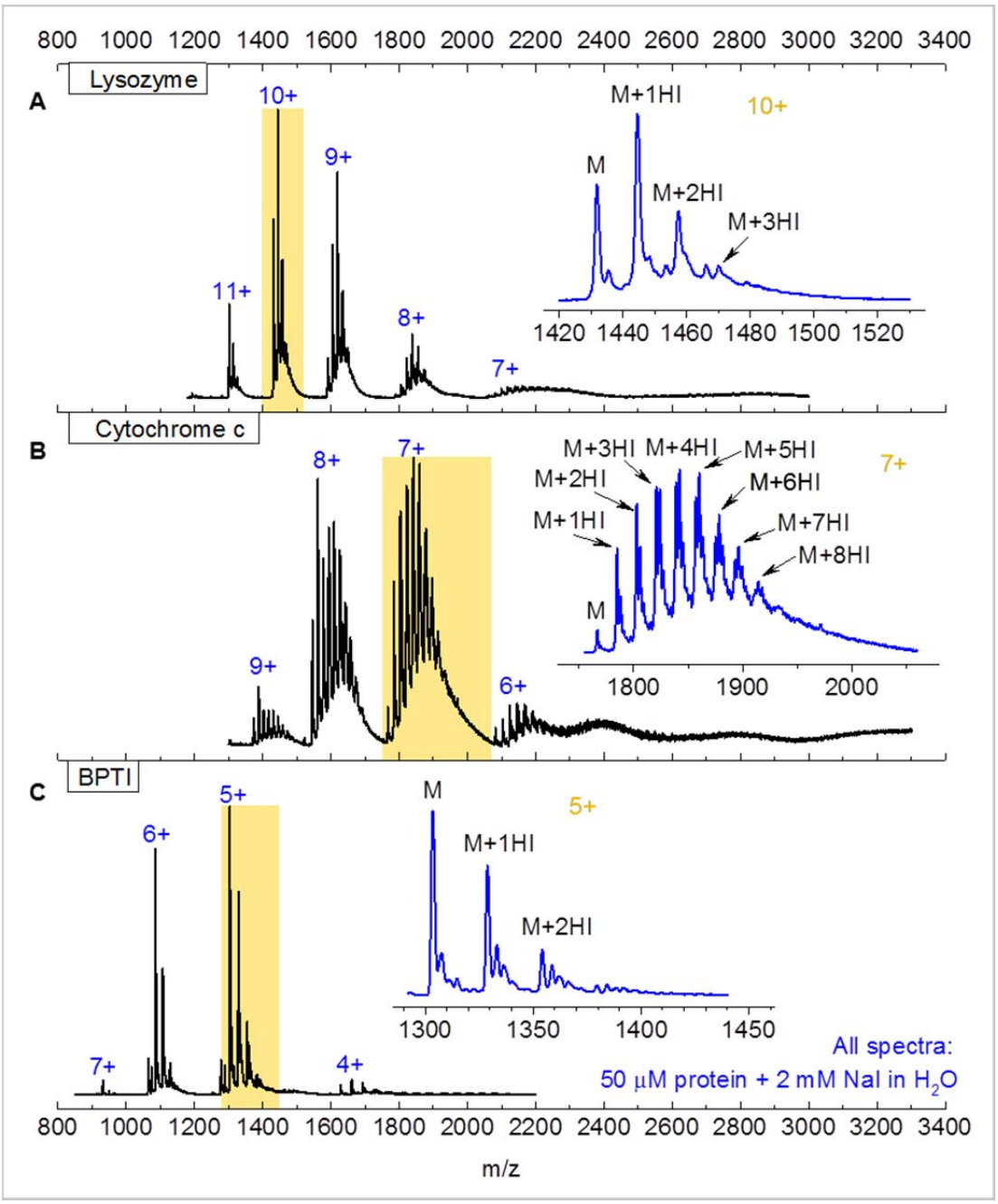


Figure 5-2 *Representative mass spectra of lysozyme (**A**), cytochrome c (**B**) and BPTI (**C**) acquired on the MoQ-ToF at 300 K in the MS-only mode. The proteins are electrosprayed in 50 μM concentrations from the aqueous solutions with the 40-fold molar excess of NaI. Charge states of the protein-salt clusters are shown in blue. Insets show zoomed-in most dominant charge states highlighted in yellow. Annotated therein are groups of peaks corresponding to protein-salt clusters, with the first peak (M) being unadducted protein, second one (M+1HI) – protein adducted with a single I^- ion, etc. Each of such clusters may or may not feature additional Na^+ ions.*

5.1.4 Validation of ion source conditions

A comparison of three different MoQ-ToF ion source temperatures was conducted to establish whether salt clustering on the proteins is affected by desolvation conditions. The relative ratios of the free and salt-bound ions were compared at different source temperatures: 60, 80 and 100 °C (the middle value being the working temperature at which all results presented here were obtained). The findings are shown in ***Figure 5-3***, and they confirm that, for all three proteins, these ratios are the same across different temperatures. Therefore, desolvation process has yielded stable gas-phase re-equilibrated populations of protein and protein-salt ions[36; 37]. This observation is consistent with previous results that proteins retain their structure during the ESI process[38] and especially its nano-spray implementation[39].

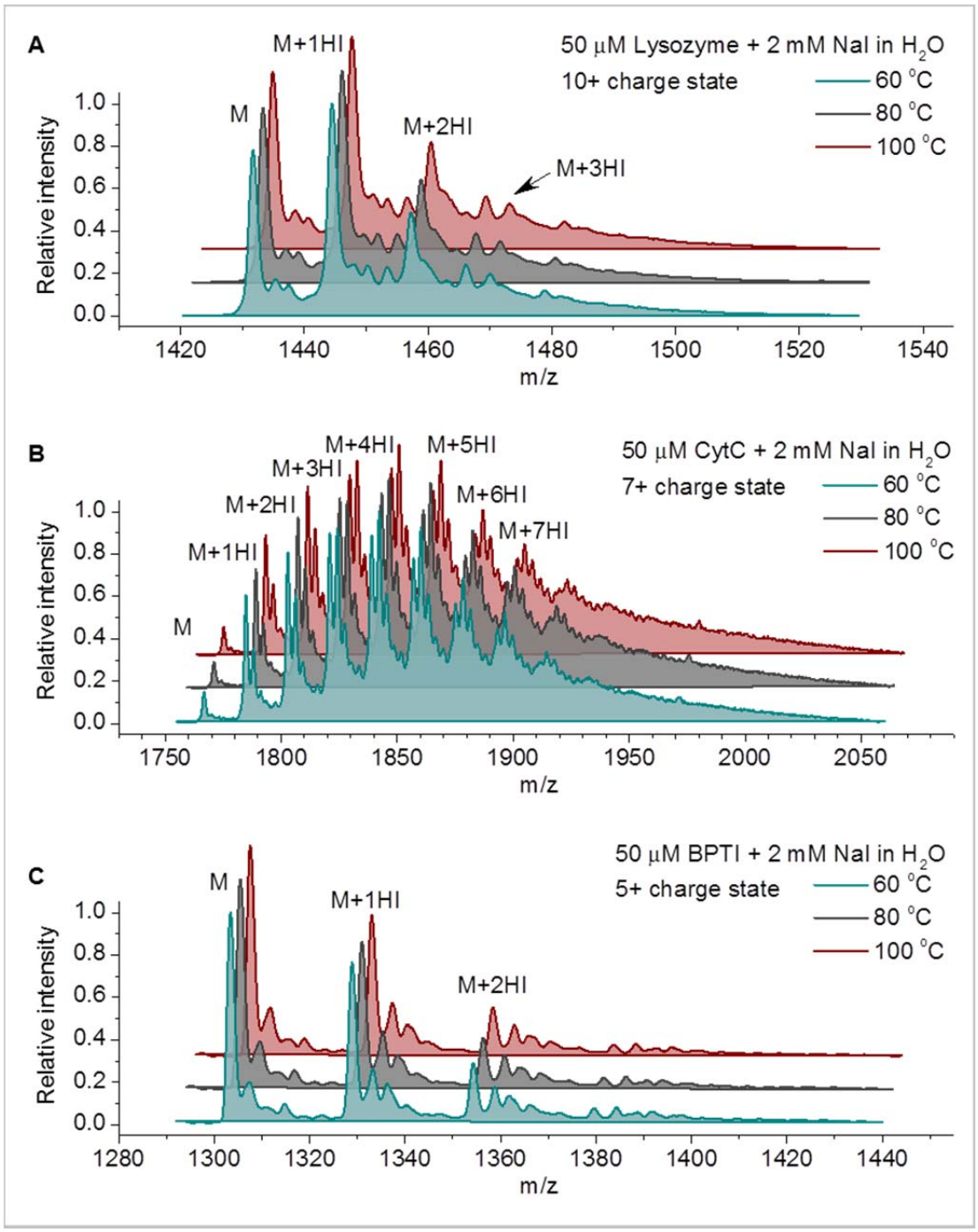


Figure 5-3 *MoQ-ToF mass spectra of the 50 μM protein (**A** – lysozyme, 10+ charge state; **B** – cytochrome c, 7+ charge state; **C** – BPTI, 5+ charge state) with 2 mM NaI in sprayed from aqueous solution at different source temperatures: 60°C (cyan), 80°C (grey) and 100°C (red). The data were acquired in the MS-only mode. Each mass spectrum is the average of 3 experimental repeats taken on different days at the same temperature.*

5.2 Lysozyme

Lysozyme-NaI clusters present themselves in 5 charge states: from 7+ to 11+, with the 10+ being most dominant (***Figure 5-2A***). The change in buffer gas temperature (especially lowering it) has a dramatic effect on salt adduction to the protein (***Figure 5-4A***), as can be seen from shifts of the protein-salt clusters. An increase in temperature causes a loss of salt ions, whereas at lower temperatures there is a higher retention of HI adducts, especially for lower charge states of the protein (7+ and 8+). This trend becomes more evident by considering the relative intensities of the non-adducted and HI-adducted lysozyme signals, and plotting the ion abundances as a function of charge state (***Figure 5-4B***). At 360 K cell temperature, the non-adducted signal is most abundant for 10+ and 11+ charge states, whereas at ambient temperature (300 K) the non-adducted signal is strongest only for the highest charge state by a small margin. At lower temperatures, the two highest charge states are dominated by singly-adducted signals. As the charge state decreases (9+ and below), more HI are retained on the protein, and this trend continues at the lowest cell temperature. These observations support the accepted view of non-covalent protein-ligand interactions becoming stronger at low temperatures[40; 41] and also that salt and, by inference, buffer salts are lost in IM-MS experiments performed at ambient temperatures.

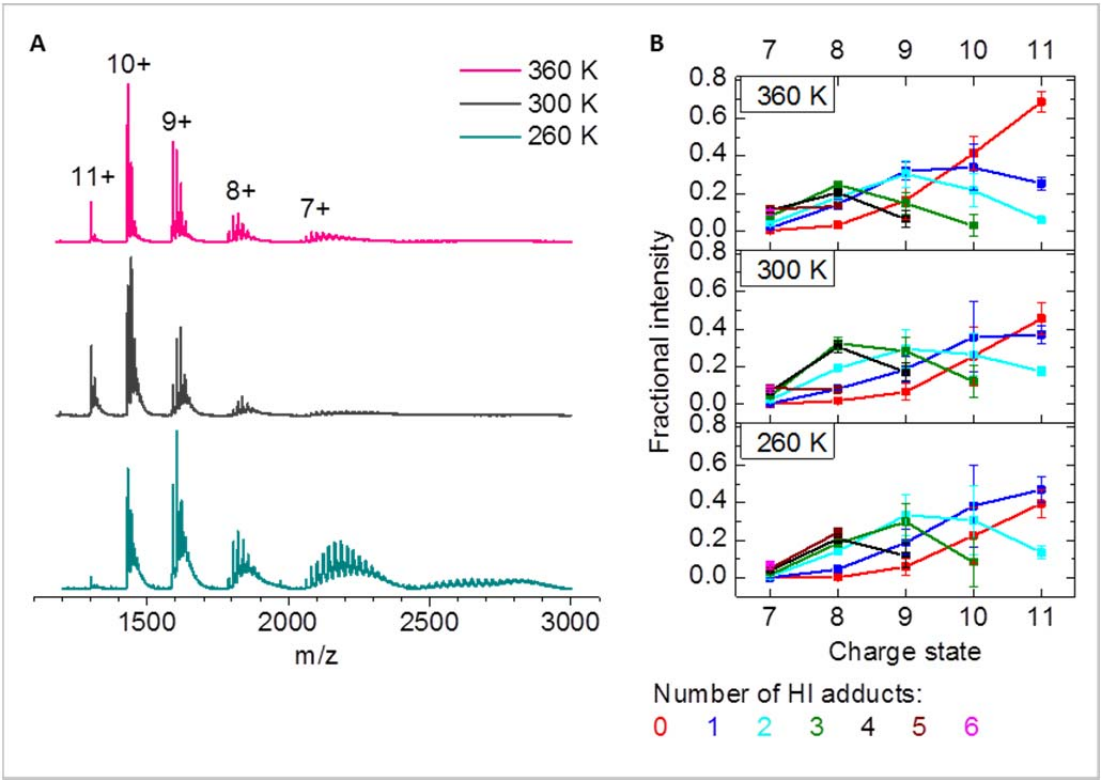


Figure 5-4 *Drift gas temperature effect on HI adduction to lysozyme. The data were acquired in the MS-only mode. Panel* ***A*** *– mass spectra of 50 μM lysozyme + 2 mM NaI in H_2O at different drift cell temperatures (260 K – turquoise, 300 K – dark grey, 360 K - pink). Panel* ***B*** *– plots of fractional ion intensity as a function of charge state at different cell temperatures for non-adducted and up to 6 HI-adducted protein ions (colour-code legend is shown below the plots). Each trace is the average of three experimental repeats conducted on different days.*

Arrival time distributions were obtained for the free protein and its single-, double- and triple-HI adducted species at three buffer gas temperatures – 260, 300 and 360 K. Results are presented as drift time distributions at drift potential 35 V, converted to CCS units and are arranged in two sets – first highlighting the effect of HI adducts on the protein (***Figure 5-5***), second – comparing the effect of different drift buffer gas temperatures (***Figure 5-7***). Both sets are comprised of same data that are re-grouped for easier viewing. Numerical values for collision cross sections were calculated for the free protein and its single-, double- and triple-HI adducted species at 300 K only. In this case however, for the 7+ ion, the salt-free ion was, albeit detectable,

too weak in intensity to provide reliable measurements on it. For the 11+ ion, the 3-adducted lysozyme was undetectable. These CCS values are plotted in ***Figure 5-8***.

5.2.1 HI effect on lysozyme conformations

Regardless of the drift cell temperature, the conformational space occupied by free and HI-adducted lysozyme shifts towards higher cross sections and widens from lower to higher charge states (***Figure 5-5***). This trend is commonly observed in ion mobility measurements and is indicative of Coulombic repulsion caused by additional charges on the protein leading to partial unfolding. Upon such unfolding, the structure becomes more flexible and its increased flexibility leads to wider conformational space available to the protein. All of the charge states display a bi-nodal drift time distribution, except for the smallest charge state, where the conformations are compacted to mainly one peak with a leading-edge shoulder. This is in good agreement with the earlier findings[42; 43; 44; 45]demonstrating that the charge states below 8+ of lysozyme are compact and possibly native-like. The 8+ charge state is therefore on the borderline of the folded and unfolded states where major conformational transitions occur.

Indeed, major differences between adducted and non-adducted ions are observed for the 8+ charge state, where the free ion features more compact populations in its earlier-arriving peak, compared to the adducted species. Addition of anions seems to cause lysozyme leave the compact sub-conformations domain resulting in tightening of the CCS distribution of the first peak, thus stabilising the fold. This effect is somewhat different at

lowered temperature (***Figure 5-5A***), where the conformations in the salt-free earlier-arriving peak are almost as compact as those in the salt-adducted peak, with significant degree of conformational inter-conversion between the two conformations taking place as a result of additional HI ions.

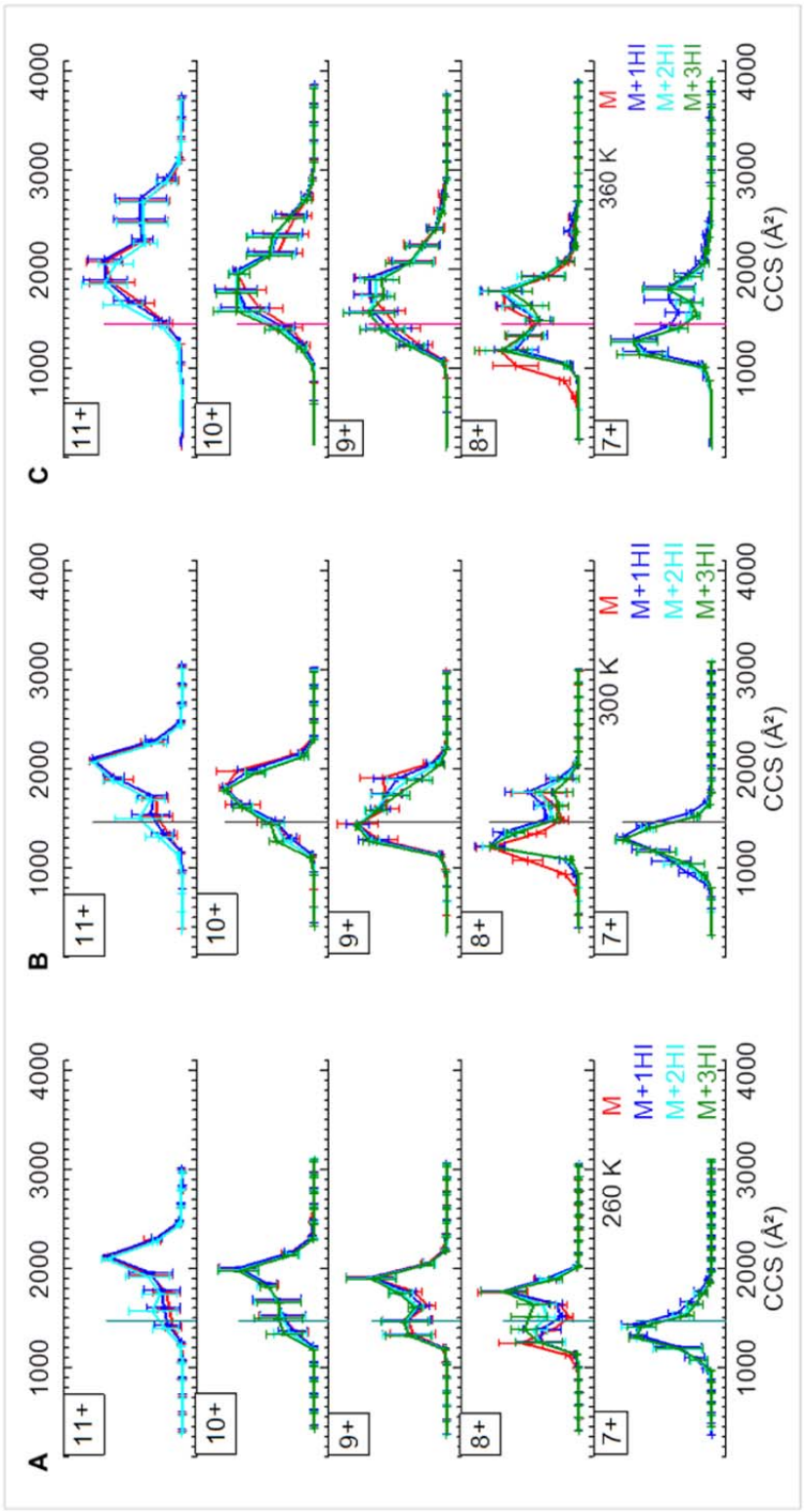


Figure 5-5 *Drift time distributions of the 7+ through 11+ charge states of lysozyme (drift time units are converted to CCS units). Measurements are carried out at the buffer gas temperatures of: 260 K (panel **A**), 300 K (panel **B**) and 360 K (panel **C**). Ion flux data for non-adducted (red) and up to 3 HI-adducted (1Γ – blue, 2Γ – cyan, 3Γ – green) protein ions are shown. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values indicated by vertical lines were calculated from their coordinates at 260 K (turquoise), 300 K (dark grey) and 360 K (pink).*

When comparing the trends caused by increasing number of HI adducts on the protein for the charge states above 8+, one can observe how salt ions 'counter-act' the Coulombic unfolding taking place due to increased charge. This effect can be observed in ***Figure 5-5***, however, to deconvolute this effect from others, selected extreme charges and adduction levels are presented in ***Figure 5-6***. Indeed, as the charge state increases, the overall bi-nodal distribution gravitates towards the larger CCS (*cf.* left and right panels in ***Figure 5-6***), whereas the more HI molecules are present on the protein, the more the balance shifts back to the smaller CCS, converging on the crystal structure values (*cf.* top and bottom panels in ***Figure 5-6***). This effect is more pronounced for higher charge states, which confirms structure-stabilising effect of chaotropes on lysozyme[25], which is one of the first proteins on which the opposite Hofmeister series behaviour was observed[46].

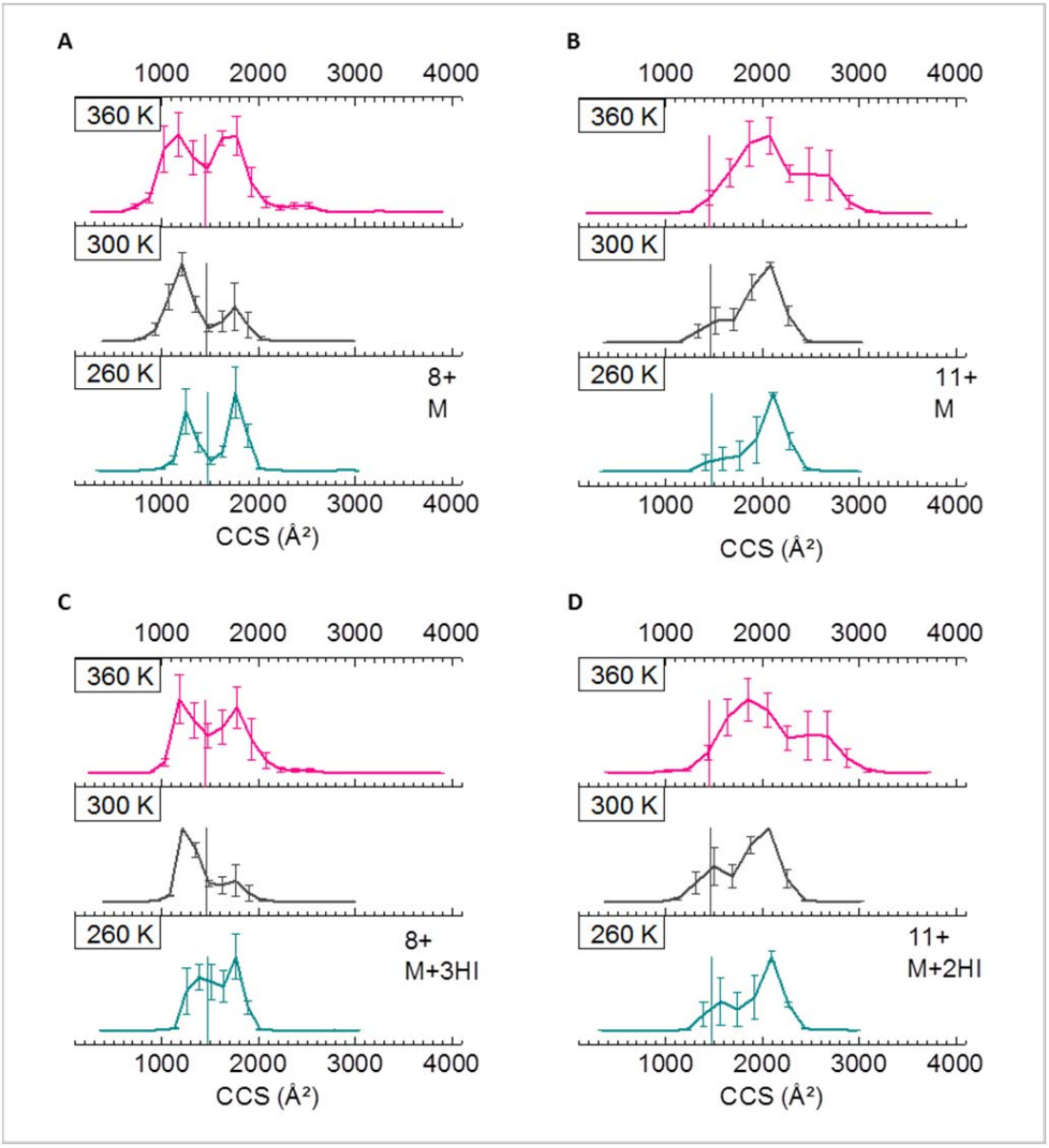


Figure 5-6 *Drift time distributions of the extreme charge states of lysozyme – 8+ (left panels: **A** and **C**) and 11+ (right panels: **B** and **D**), and extreme adduction level – apo (top panels: **A** and **B**) and highest adduction measured (bottom panels: **C** and **D**). Measurements are carried out at the buffer gas temperatures of: 260 K (turquoise), 300 K (dark grey) and 360 K (pink). Each trace is the average of three experimental repeats conducted on different days. Crystal structure values are indicated by the vertical lines (calculated from their coordinates with MobCal for each of the temperatures).*

5.2.2 Drift gas temperature effect on lysozyme conformations

Lysozyme, in its all states of adduction including free form, responds to the change in the drift gas temperature in a uniform manner, as is evident when comparing the data arranged in each of the horizontal rows (***Figure 5-7***). At lowered buffer gas temperatures (260 K), two conformers are better resolved, which is expected at lower temperatures[47; 48]. Also, the peaks are narrower, indicating conformational stabilisation and decreased inter-conversion due to 'freezing'. A striking, and seemingly counter-intuitive, feature of the protein's behaviour at low temperature is that larger conformations are favoured compared to ambient temperature (300 K). Indeed, the 7+ ion drift profile indicated one peak, albeit with a small shoulder, whereas as early as the 8+ and all charge states above, this very peak becomes minor as the later-arriving peak takes dominance. This phenomenon is known as 'cold denaturation'[49; 50] and is a result of a temperature-induced change in the hydrogen bonding that stabilises α-helices and β-sheets in the protein. This leads to the loss of hydrophobicity of such α-helices and destabilisation of β-sheets, resulting in denaturation[51]. For lysozyme, cold denaturation was previously observed using pressure-assisted NMR spectroscopy studies[52]. The crystal structure CCS values fall exactly between the two conformational groups observed in the experiment at the ambient- and low-temperature drift gas for the 8+ charge state. For the lower charge state (7+), the experimental values are mainly just below it, and for all charge states above (9+ through 11+) the less-abundant conformational population has similar CCS values to those of the crystal structure, and the higher-abundant species have noticeably higher values.

At higher drift gas temperatures (360 K), a distinct bi-nodal drift profile is observed for the 7+ charge state (*cf*. same charge state at 300 and 260 K), suggesting thermal melting onset with the majority of the conformations lying below the crystal structure value and about a one-third of them above that value. For the 8+ charge state, both populations are equally represented, and from the 9+ and up, the major peak is found above the crystal structure, with only a minor fronting shoulder covering those values. The 10+ and 11+ charge state feature an additional distinct peak above 2500 $Å^2$ due to further protein unfolding. It is noteworthy that increasing number of HI on these charge states does somewhat increase the abundance of the more compact population (horizontal rows of 10+ and 11+), thus 'resisting' the thermal unfolding of lysozyme.

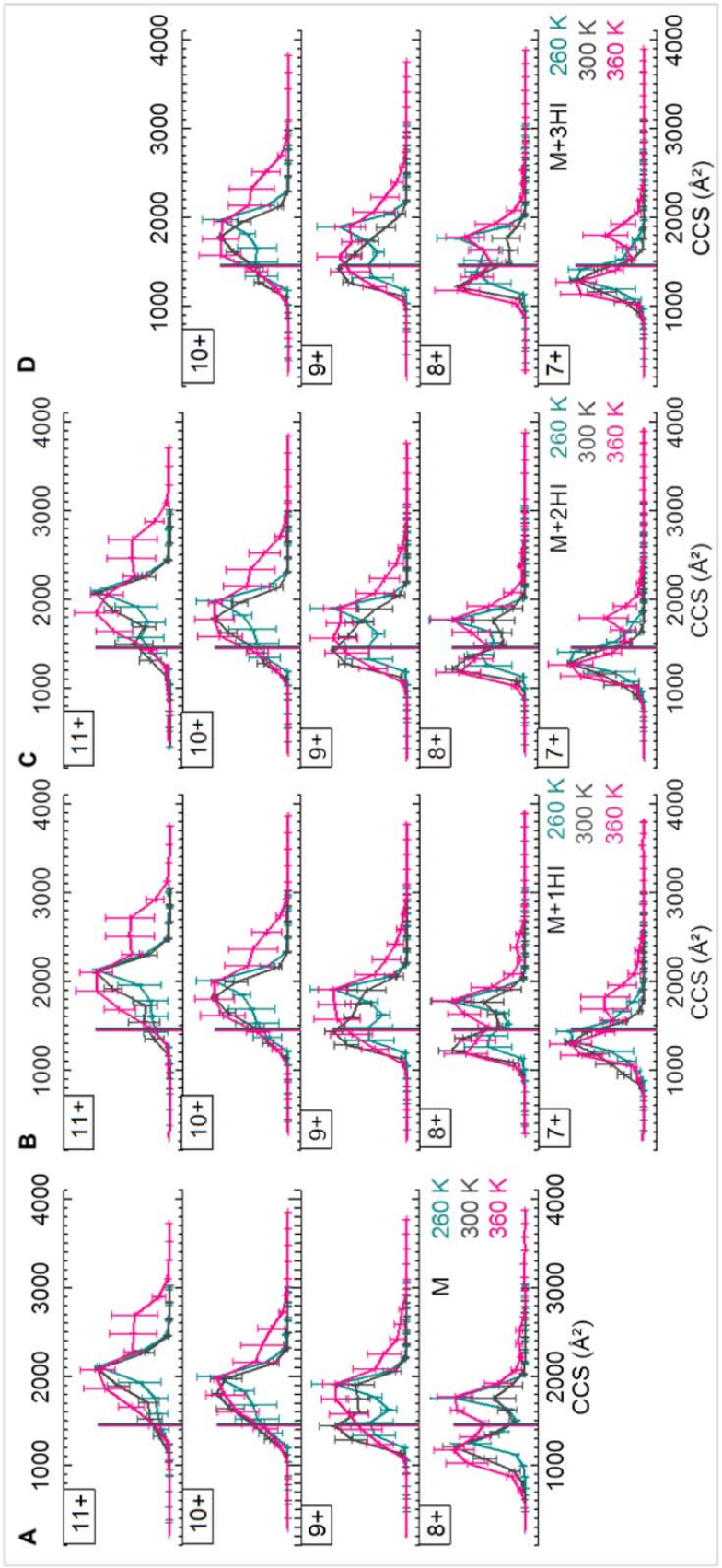


Figure 5-7 *Drift time distributions of the 7+ through 11+ charge states of lysozyme (drift time units are converted to CCS units). Measurements are carried out at the buffer gas temperatures of: 260 K (turquoise), 300 K (dark grey) and 360 K (pink). Ion flux data for non-adducted (panel **A**) and up to 3 HI-adducted (1Γ – panel **B**, 2Γ – panel **C**, 3Γ – panel **D**) protein ions are shown. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values indicated by vertical lines were calculated from their coordinates at 260 K (turquoise), 300 K (dark grey) and 360 K (pink) (N.B. these values a very close to each other and may look line one vertical line).*

5.2.3 Collision cross sections of free and HI-adducted lysozyme at 300 K

Figure 5-8 shows the CCS values of the drift peak maxima for each observed charge state that have been plotted as a function of adduction level, at 300 K. The 'distance' between the values of two conformers for each drift profile has been shaded to assist data viewing. The crystal structure CCS value was obtained following a gas-phase minimisation as described in ***section 2.7.3*** and is shown in dashed line in ***Figure 5-8***. This value falls in the centre of CCS values distribution for the 7+ and 8+ ions suggesting their compact crystal structure-like packing as previously shown by Clemmer group[44] and ours[53]. All higher-charged ions are above the crystal structure CCS, however each additional HI on the protein assists its compaction bringing its size closer to the crystal structure value. Somewhat anomalous behaviour is observed for the 8+ ion, where addition of the first HI results in a CCS increase; all subsequent salt ions conform with the overall trend of reducing the CCS space that the protein is in.

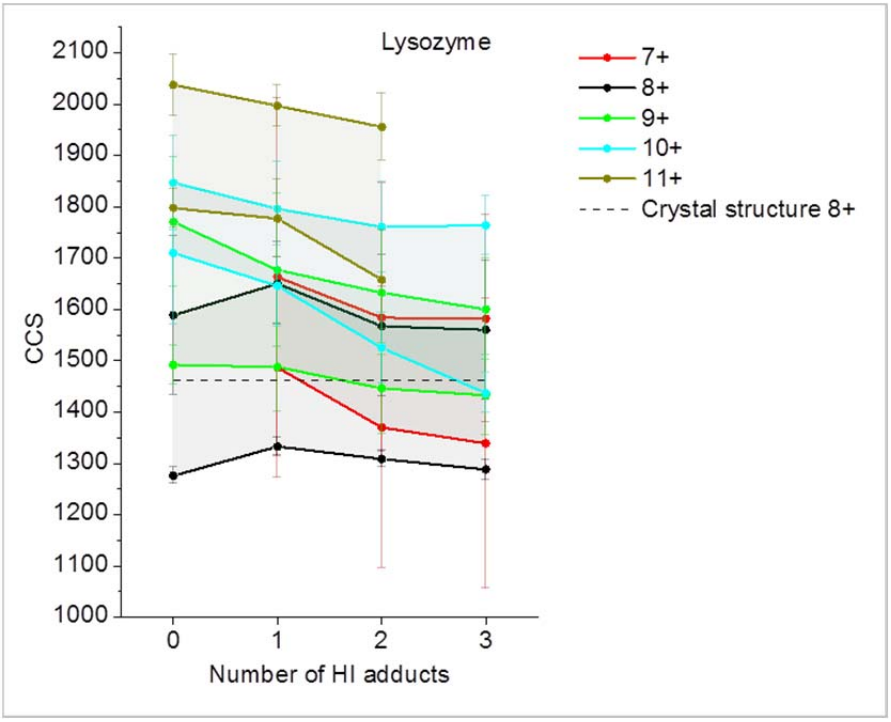


Figure 5-8 *Summary of cross sections determined for all ions observed for lysozyme at 300 K buffer gas temperature. Data-points indicate drift peak maxima for each observed charge state that have been plotted as a function of adduction level. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values for the apo-protein at 300 K are indicated by the dashed horizontal line. Numerical CCS values for both experimental and calculated CCS values for each of the data points are tabulated in* ***Appendix, Table A 3*** *and* ***Table A 4****.*

5.3 Cytochrome *c*

Cytochrome *c*-NaI clusters present themselves in 4 charge states: from 6+ to 9+, with the 7+ and 8+ being equally most dominant (***Figure 5-9A***). The change in buffer gas temperature (especially lowering it) has a dramatic effect on salt adduction to the protein (***Figure 5-9A***), as can be seen form shifts of the protein-salt clusters. In contrast to lysozyme, cytochrome *c* does not exhibit a significant loss of salt ions at increased temperature; moreover an opposite effect is observed especially for the lowest charge state. At lower

temperatures a dramatic increase in the levels of HI adduction is observed, again, especially for lowest charge states of the protein (6+). These trends become more evident by measuring the relative intensities of the non-adducted and HI-adducted cytochrome *c* signals on the mass spectra, and plotting the ion abundances as a function of charge state (***Figure 5-9B***). At 360 K cell temperature, the non-adducted signal, across all charge states, is very similar to the one at 300 K. All adducted species become slightly more prominent at higher temperature for all charge states except the lowest. The 6+ charge state exhibits a different trend: high levels of adduction are not favourable at 360 K. At lower temperature, the non-adducted signal is almost lost for all charge states, and higher adduction numbers are favoured. Interestingly, the adduction number distribution is very tight for the 6+ charge state.

Overall for cytochrome *c*, the number of adducts is similar across all charge states except for the lowest within one temperature setting, unlike for lysozyme, where adduction number is highly charge-state dependent. Hence, the adduction level of the 6+ charge state is the one affected by the temperature most of all.

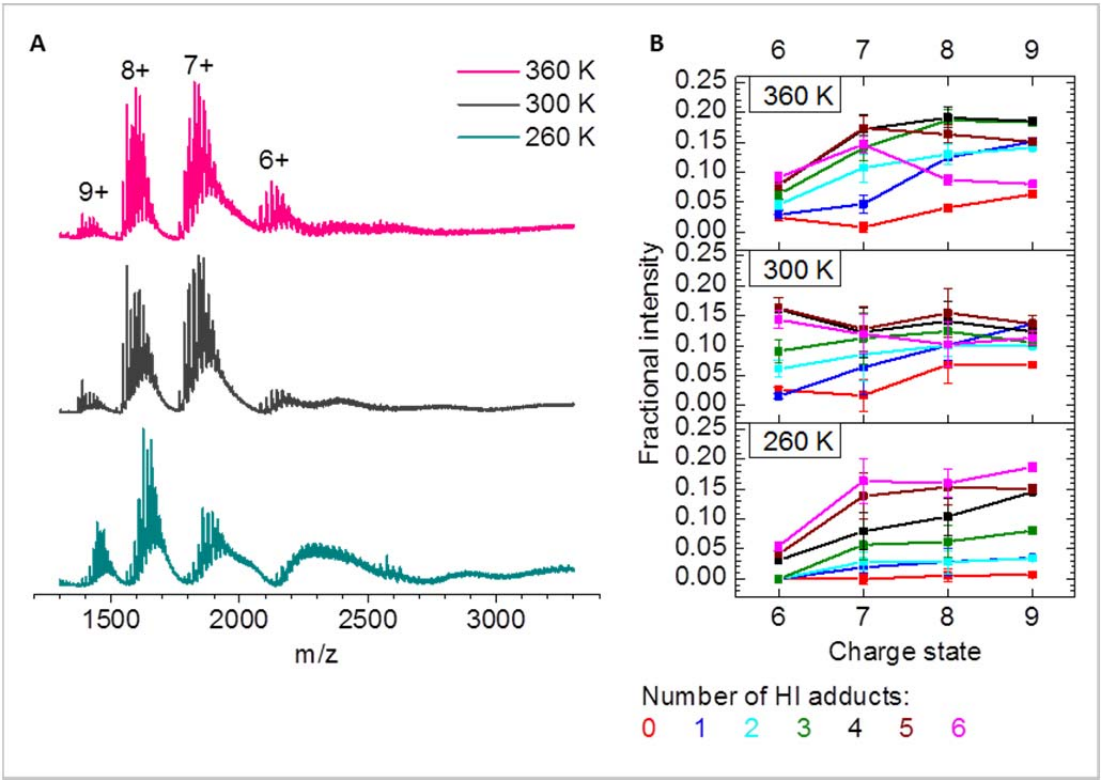


Figure 5-9 *Drift gas temperature effect on HI adduction to cytochrome c. The data were acquired in the MS-only mode. Panel* ***A*** *– mass spectra of 50 μM cytochrome c + 2 mM NaI in H_2O at different drift cell temperatures (260 K – turquoise, 300 K – dark grey, 360 K – pink). Panel* ***B*** *– plots of fractional ion intensity as a function of charge state at different cell temperatures for non-adducted and up to 6 HI-adducted protein ions (colour-code legend is shown below the plots). Each trace is the average of three experimental repeats conducted on different days.*

Arrival time distributions were obtained for the free protein and its single-, double- and triple-HI adducted species at three buffer gas temperatures – 260, 300 and 360 K. For the 7+ ion at 360 K the non-adducted cytochrome *c* was undetectable. Results are presented as drift time distributions at drift potential 35 V, converted to CCS units and are arranged in two sets – first highlighting the effect of HI adducts on the protein (***Figure 5-10***), second – comparing the effect of different drift buffer gas temperatures (***Figure 5-11***). Both sets are comprised of same data that are re-grouped for easier viewing. Numerical values for collision cross sections were calculated for the free protein and its single-, double- and triple-HI adducted species at 300 K only.

For the extreme (6+ and 9+) ions, however, the doubly- and triply-adducted ions were, albeit detectable, very weak in intensity to provide reliable measurements on them. These CCS values are plotted in ***Figure 5-12***.

5.3.1 HI effect on cytochrome c conformations

Regardless of the drift cell temperature, the conformational space occupied by free and HI-adducted cytochrome *c* shifts towards higher cross sections and widens from lower to higher charge states (***Figure 5-10***). This is the manifestation of Coulombic repulsion effect that is discussed above for lysozyme. All of the charge states display a bi-nodal drift time distribution, except for the 6+ at ambient temperature and 7+ at increased temperature, where multiple conformations inter-convert and elute as one tailing peak. At ambient and increased temperature, majority of conformations lies within the earlier-arriving peak and below the calculated crystal structure values (***Figure 5-10B*** and ***C***); at low temperature, the effect is opposite (***Figure 5-10A***).

Major differences between adducted and non-adducted ions are observed for the earlier-arriving peak at 260 K; the free ion populates a wider range of conformations in its earlier-arriving peak, compared to the adducted species (***Figure 5-10A***). Addition of anions increases the population of the compact conformational node resulting in tightening of the CCS distribution of the first peak, presumably stabilising the fold. At ambient and increased temperatures, the ATD profiles do not change significantly with HI addition (***Figure 5-10B*** and ***C***), with an exception for z = 9, where the singly-adducted protein has a slight preference for more unfolded conformations. Overall, HI

addition has a very limited effect on cytochrome *c* conformations, except a minor 'tightening' of the smaller conformers at lower temperature.

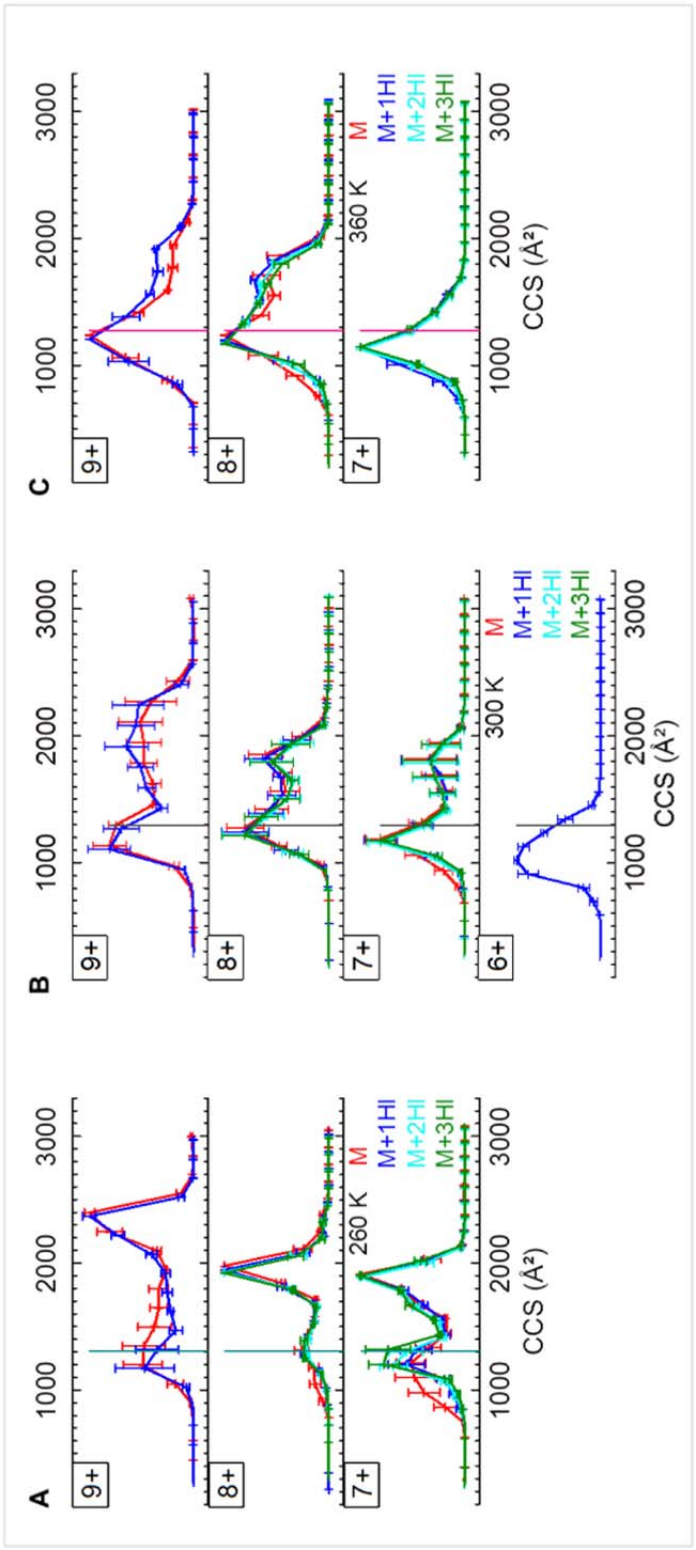


Figure 5-10 *Drift time distributions of the 6+ through 9+ charge states of cytochrome c (drift time units are converted to CCS units). Measurements are carried out at the buffer gas temperatures of: 260 K (panel **A**), 300 K (panel **B**) and 360 K (panel **C**). Ion flux data for non-adducted (red) and up to 3 HI-adducted (1Γ – blue, 2Γ – cyan, 3Γ – green) protein ions are shown. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values indicated by vertical lines were calculated from their coordinates at 260 K (turquoise), 300 K (dark grey) and 360 K (pink).*

5.3.2 Drift gas temperature effect on cytochrome c conformations

Unlike with lysozyme, lowered buffer gas temperatures (260 K), do not seem to improve the resolution of the two conformers (***Figure 5-11***). Similarly to lysozyme, cytochrome *c*, in its all states of adduction including the free form, responds to the change in the drift gas temperature in a uniform manner, as is evident when comparing the data arranged in each of the horizontal rows (***Figure 5-11***). The 'cold denaturation' phenomenon holds true for cytochrome *c*, as well as for lysozyme. For cytochrome *c*, cold denaturation was previously observed using CD spectroscopy studies[54]. A striking difference between cytochrome *c* and lysozyme is observed at increased cell temperatures (360 K). Here, heating causes a decrease in the CCS, bringing the majority of the conformational populations just under the values of the crystal structure. The effect observed here is probably caused by thermal collapse of the protein fold. This observation is contrary to the earlier findings by Mao *et al.*[55], however it must be mentioned that higher ion injection energies were used in that work which might cause protein unfolding. The ambient temperature ATD profiles are 'in transition' between the two extreme temperatures, and are closer to the ATDs at 360 K than 260 K.

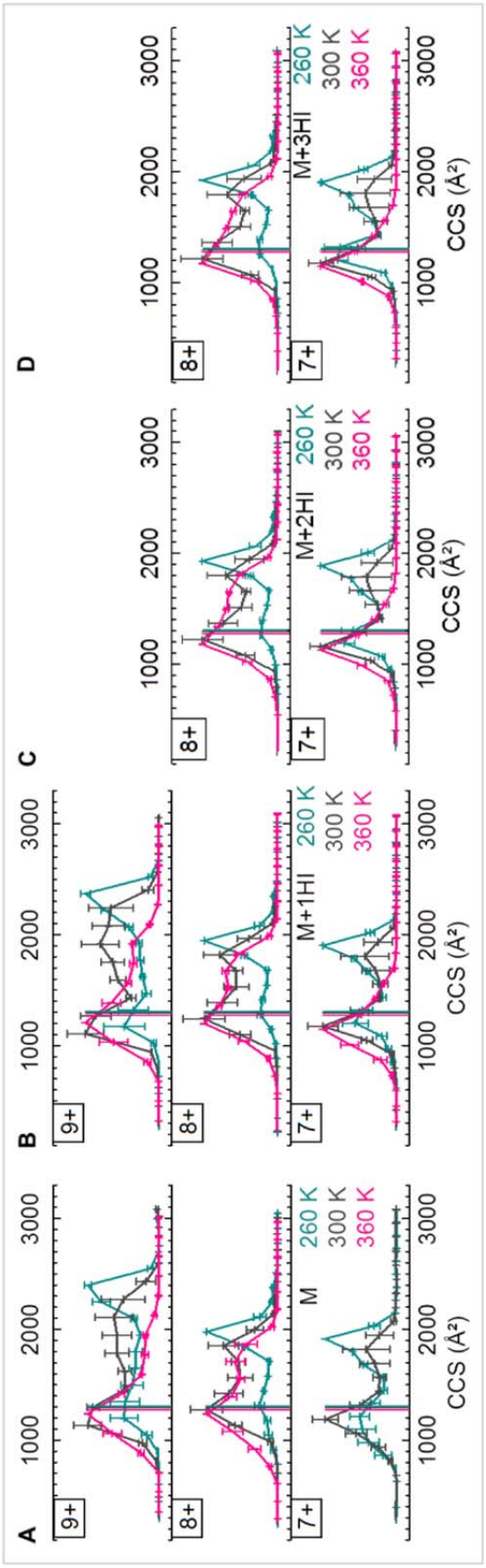


Figure 5-11 *Drift time distributions of the 7+ through 9+ charge states of cytochrome c (drift time units are converted to CCS units). Measurements are carried out at the buffer gas temperatures of: 260 K (turquoise), 300 K (dark grey) and 360 K (pink). Ion flux data for non-adducted (panel **A**) and up to 3 HI-adducted (1Γ – panel **B**, 2Γ – panel **C**, 3Γ – panel **D**) protein ions are shown. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values indicated by vertical lines were calculated from their coordinates at 260 K (turquoise), 300 K (dark grey) and 360 K (pink) (N.B. these values a very close to each other and may look line one vertical line).*

5.3.3 Collision cross sections of free and HI-adducted cytochrome c at 300 K

Figure 5-12 shows the CCS values of the drift peak maxima for each observed charge state plotted as a function of adduction level, at 300 K. The 'distance' between the values of two conformers for each drift profile has been shaded to assist data viewing. The crystal structure CCS value was obtained following a gas-phase minimisation as described in ***section 2.7.3*** and is shown in dashed line in ***Figure 5-12***. A trend similar to that observed for lysozyme is noted for cytochrome *c*: the CCS value from the crystal structure falls in the lower part of the CCS values distribution for the 7+ ions suggesting their compact packing. The CCS values of the 6+ ions lie below those of the crystal structure, and the ions with charge states above 7+ have higher CCS than that of the crystal structure. This is in good agreement with earlier work by Jarrold group[56] and ours[57] that established that the 7+ charge state is on the border line of compact and extended conformations, with folded structures only being stable below 7+ charge state. The higher the charge state is the fewer HI are required to cause compaction of the conformational distribution.

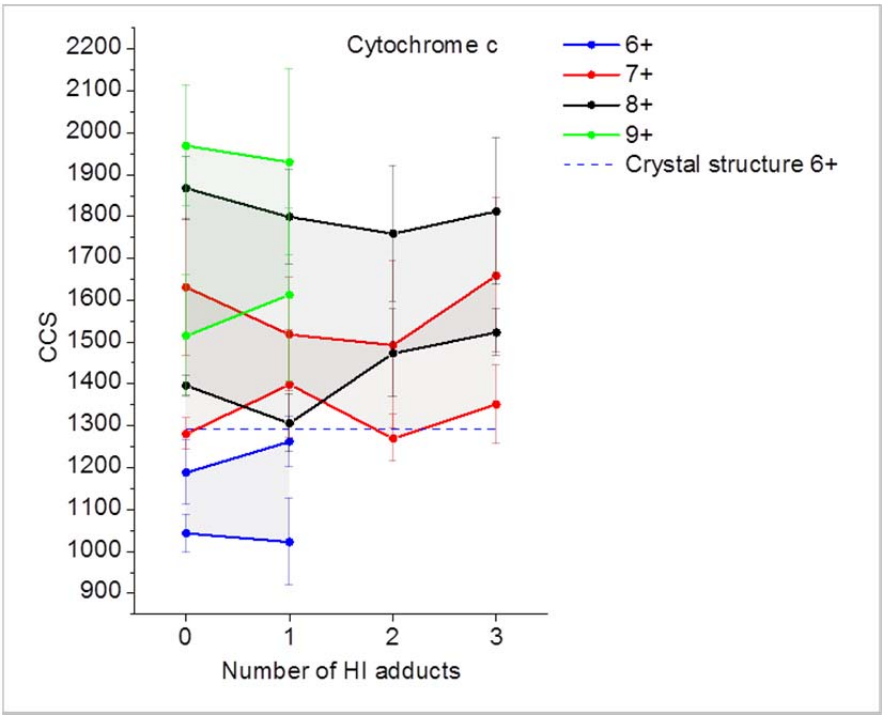


Figure 5-12 *Summary of cross sections determined for all ions observed for cytochrome c at 300 K buffer gas temperature. Data-points indicate drift peak maxima for each observed charge state that have been plotted as a function of adduction level. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values for the apo-protein at 300 K are indicated by the dashed horizontal line. Numerical CCS values for both experimental and calculated CCS values for each of the data points are tabulated in* ***Appendix****,* ***Table A 3*** *and* ***Table A 4****.*

5.4 BPTI

BPTI-NaI clusters present themselves in 4 charge states: from 4+ to 7+, with the 5+ and 6+ species being dominant (***Figure 5-13A***). Similarly to the two previously discussed proteins, the most dramatic effect on the mass spectra is caused by lowering temperature rather than raising it. This is evidenced by increased adduction number on the two lowest charge states (4+ and 5+). In contrast to the both larger proteins described above, an increase in temperature causes an increase in the level of observed adduction in salt ions

and the two extreme charge states (4+ and 7+) are more populated than at 260 K and 300 K. Plots of the relative intensities of the non-adducted and HI-adducted BPTI signals on the mass spectra *vs* charge state (***Figure 5-4B***) show the HI adduction behaviour at all temperatures considered during the experiment. A significant difference is observed for the 360 K, where higher charge states retain HI adducts better than at two other temperatures. The non-adducted BPTI has the strongest signal at ambient temperature. The singly-adducted protein is the favoured arrangement of all HI-BPTI interactions across all temperatures, but only marginally so at 360K. This suggests that the population of the adducted species in the mass spectrometer alters during the experiment. If we are to consider the data taken at 260 K as a baseline (since we cannot be adding salt in the drift cell!) then we observe that the mobility experiment at 300 K causes a decrease in the high *m*/*z* poorly resolved salted forms of the protein, and this has populated the apo forms most. By contrast, the mobility experiment at 360 K has caused a different redistribution of these salty peaks populating more adducted forms, especially for the z = 6 and 7 species. We can speculate that the protein is more flexible at the higher temperature and that more favourable binding sites are accessible to the HI.

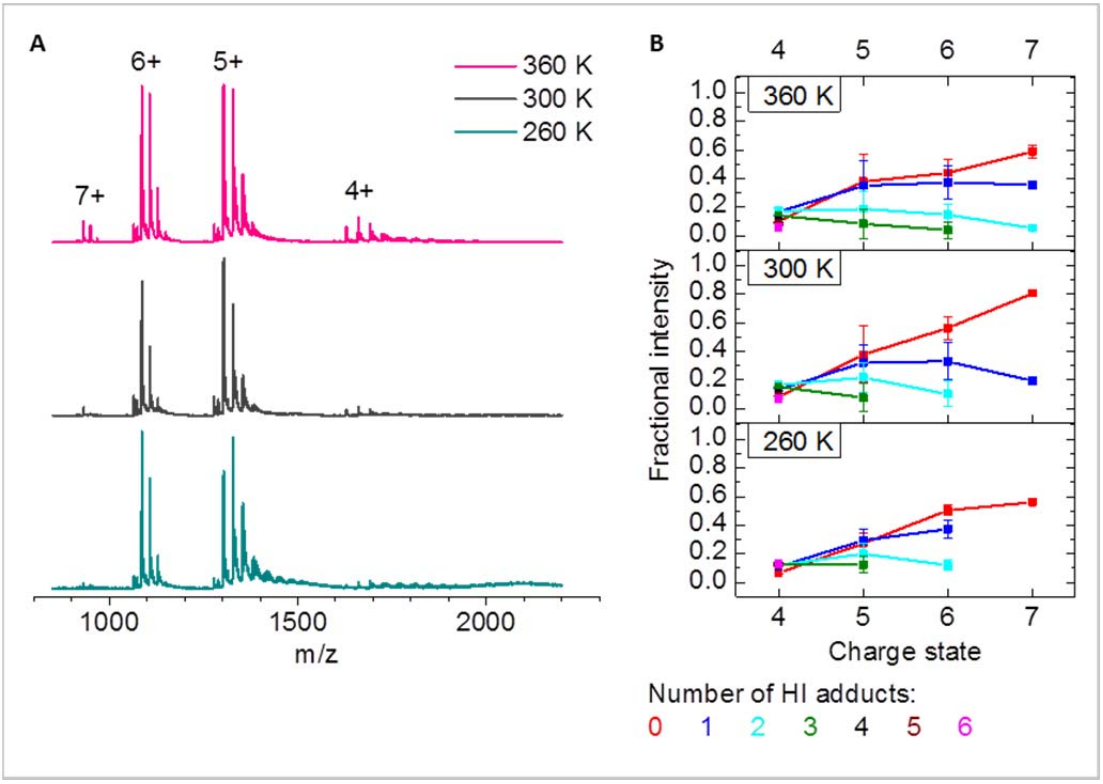


Figure 5-13 *Drift gas temperature effect on HI adduction to BPTI. The data were acquired in the MS-only mode. Panel* ***A*** *– mass spectra of 50 μM BPTI + 2 mM NaI in H_2O at different drift cell temperatures (260 K – turquoise, 300 K – dark grey, 360 K - pink). Panel* ***B*** *– plots of fractional ion intensity as a function of charge state at different cell temperatures for non-adducted and up to 6 HI-adducted protein ions (colour-code legend is shown below the plots). Each trace is the average of three experimental repeats conducted on different days.*

Arrival time distributions were obtained for the free protein and its single-, double- and triple-HI adducted species at three buffer gas temperatures – 260, 300 and 360 K. Higher adduction levels are very low in abundance to obtain reliable values for all 6+ and 7+ ions at 360 K. Results are presented as drift time distributions at drift potential 35 V, converted to CCS units and are arranged in two sets – the first highlighting the effect of HI adducts on the protein (***Figure 5-14***), the second – comparing the effect of different drift buffer gas temperatures (***Figure 5-15***). Both sets are comprised of same data that are re-grouped for easier viewing. Numerical values for collision cross sections were calculated for the free protein and its single- and double-HI adducted species at 300 K only, wherever the signals of required intensity

were obtained. Higher adduction levels are very low in abundance to obtain reliable values. These CCS values are plotted in ***Figure 5-16***.

5.4.1 HI effect on BPTI conformations

Similarly to both proteins discussed above the conformational space occupied by free and HI-adducted BPTI shifts towards higher cross-sections and widens from lower to higher charge states, regardless of the drift cell temperature (***Figure 5-14***). Only the lowest charge state (4+) displays a bi-nodal drift time distribution, with a low-abundance fronting peak. At 300 K and largely at 360 K, the 7+ ion falls well into the crystal structure values. At 260 K, this is the case for the 6+. HI adduction does not have an effect on the ATD profiles for the 4+ and 5+ charged ions, except for the earlier-arriving conformations of the 4+ at ambient temperature (***Figure 5-14B***). Here, HI molecules tend to decrease the abundance of this highly unstable fold (its instability is evidenced by large error bars on the plot of non-adducted BPTI). All higher charged ions (6+ and 7+) experience decrease in CCS values when a single HI is attached to the protein (see two top rows in ***Figure 5-14***), and are brought closer to the crystal structure value. Interestingly, the ATD of the apo 7+ ion at ambient temperature falls into the crystal structure CCS, whereas at lower temperature it is somewhat above it, and at higher temperature it is below it (by a very small margin). Evidently, cold denaturation is observed for the highest charge state of BPTI only, this may be attributed to the presence of the stabilising disulphide bridges. For BPTI, cold denaturation was previously observed using magnetic relaxation dispersion studies in solution[58].

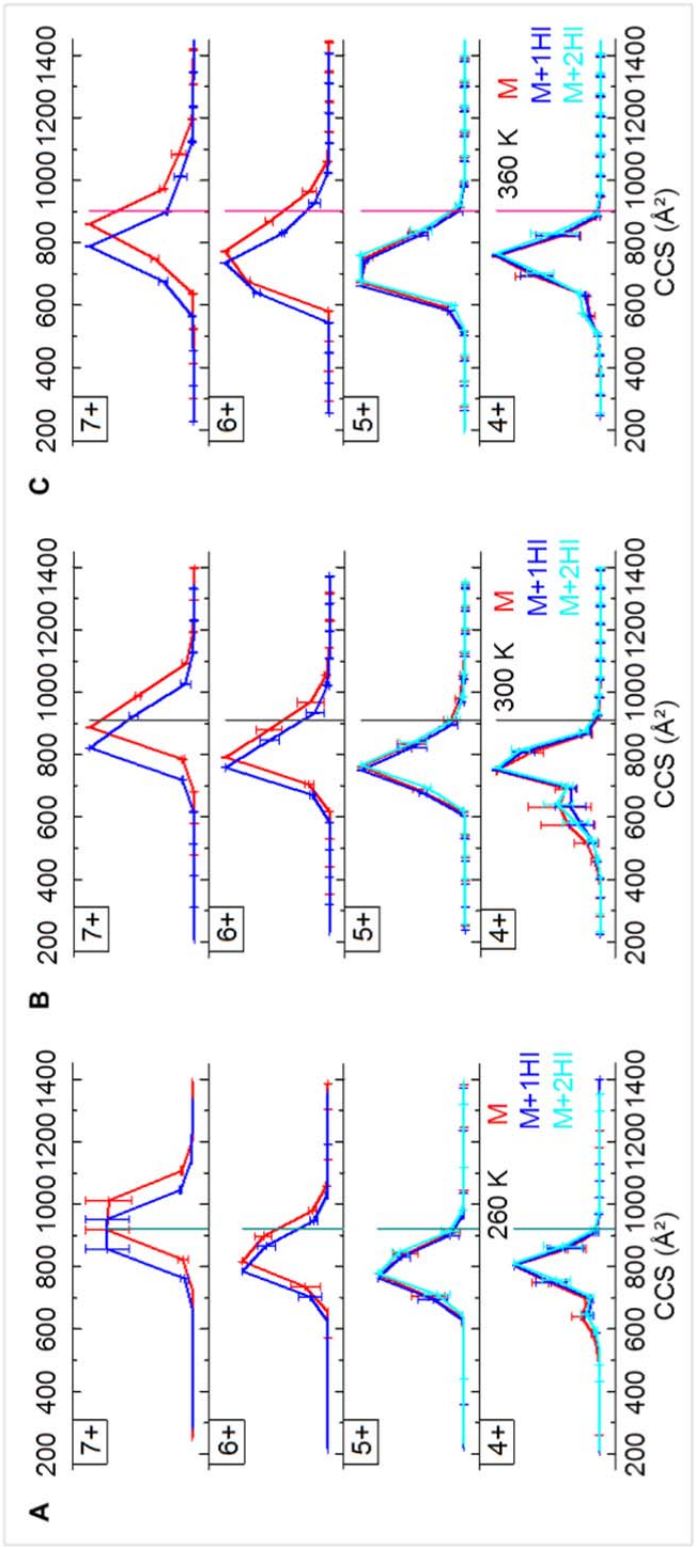


Figure 5-14 *Drift time distributions of the 7+ through 11+ charge states of BPTI (drift time units are converted to CCS units). Measurements are carried out at the buffer gas temperatures of: 260 K (panel **A**), 300 K (panel **B**) and 360 K (panel **C**). Ion flux data for non-adducted (red) and up to 2 HI-adducted (1Γ – blue, 2Γ – cyan) protein ions are shown. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values indicated by vertical lines were calculated from their coordinates at 260 K (turquoise), 300 K (dark grey) and 360 K (pink).*

5.4.2 Drift gas temperature effect on BPTI conformations

Similarly to both proteins discussed above, BPTI, in its all states of adduction including free form, responds to the change in the drift gas temperature in a uniform manner, as is evident when comparing the data arranged in each of the horizontal rows (***Figure 5-15***). Moreover, the ATD profiles are almost identical across different adduction levels (*cf.* panels ***A***, ***B*** and ***C*** in ***Figure 5-15***). A bi-nodal distribution of the ion flux is observed for the lowest charge state only (4+), featuring a small fronting peak. This peak has higher abundance at ambient temperature and represents a conformationally unstable population of ions, as evidenced by large error bars. This population is almost extinct at both lower and higher temperatures. A marginal cold denaturation effect, as evidenced by increase in CCS, is observed at lower temperatures across all adduction levels, with the effect more profound for the extreme charge states (4+ and 7+). A significant collapse of the protein fold takes place as a result of buffer gas heating for 5+ and above. For the 4+ ion, this effect is less profound, and mainly leads to the loss of the earlier-arriving peak. The crystal structure CCS values are in good agreement with the CCS of the 7+ ions at 260 K for non-adducted ion and at 300 K for singly-adducted BPTI. All other charge states feature more compact conformations.

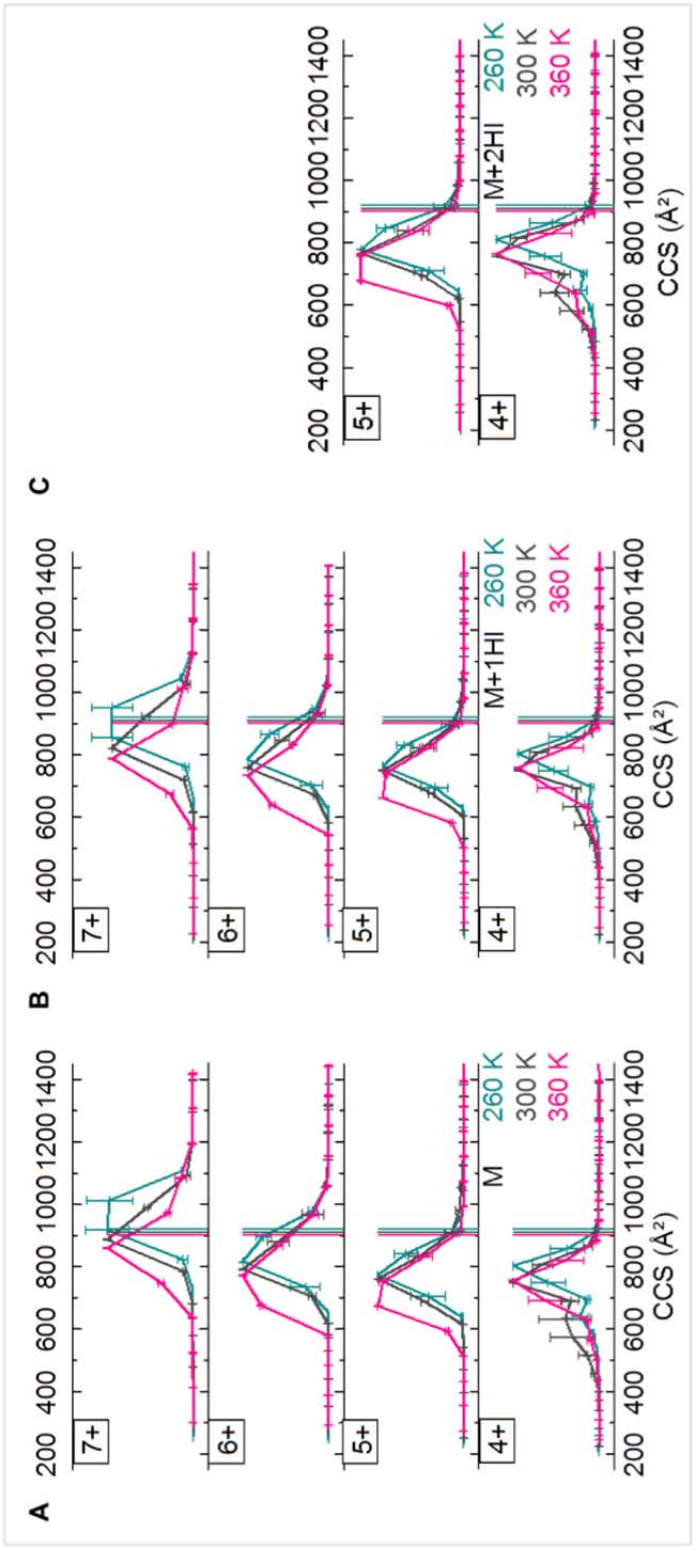


Figure 5-15 *Drift time distributions of the 7+ through 11+ charge states of BPTI (drift time units are converted to CCS units). Measurements are carried out at the buffer gas temperatures of: 260 K (turquoise), 300 K (dark grey) and 360 K (pink). Ion flux data for non-adducted (panel **A**) and up to 2 HI-adducted (1Γ – panel **B**, 2Γ – panel **C**) protein ions are shown. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values indicated by vertical lines were calculated from their coordinates at 260 K (turquoise), 300 K (dark grey) and 360 K (pink) (N.B. these values a very close to each other and may look line one vertical line).*

5.4.3 Collision cross sections of free and HI-adducted BPTI at 300 K

Figure 5-16 shows the CCS values of the drift peak maxima for each observed charge state that have been plotted as a function of adduction level, at 300 K. The 'distance' between the values of two conformers for each drift profile has been shaded to assist data viewing. The crystal structure CCS value was obtained following a gas-phase minimisation as described in ***section 2.7.3*** and is shown in dashed line in ***Figure 5-16***. This value is just below the highest experimentally measured CCS of the unadducted 7+ ion indicating that the gas-phase structure of BPTI is very compact and the 6+, 5+ and 4+ are below that value, listed in the order of decreasing CCS. This trend agrees well with that previously reported by Shelimov *et al.*[59], however it must be pointed out that these authors quoted single values for each of the charge states, whereas here the 'spread' of values between two conformational populations is given. The two highest (7+ and 6+) charge states respond to a single HI adduction in significant tightening of the fold. A similar effect is observed for the 5+ ion, with the second HI adduction brining no change in conformation. The 4+ ion, exhibits the opposite effect: no change in CCS on addition of the first HI, and decrease in CCS as a result of the second salt attachment.

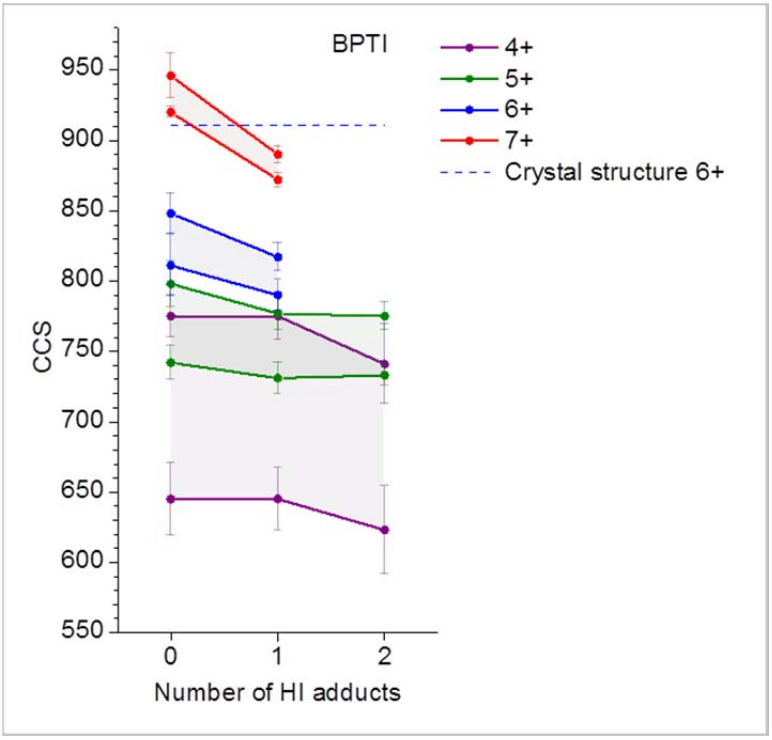


Figure 5-16 *Summary of cross sections determined for all ions observed for BPTI at 300 K buffer gas temperature. Data-points indicate drift peak maxima for each observed charge state that have been plotted as a function of adduction level. Each trace is the average of three experimental repeats conducted on different days. Crystal structure values for the apo-protein at 300 K are indicated by the dashed horizontal line. Numerical CCS values for both experimental and calculated CCS values for each of the data points are tabulated in* ***Appendix***, ***Table A 3*** *and* ***Table A 4****.*

5.5 Conclusions

Discussed here are gas-phase study results for protein ions formed out of solutions containing iodide salts. Presented findings draw a somewhat different picture from the observations of Merenbloom *et al.*[25], and suggest that the presence of the salt in the same solution as the protein is indeed sufficient to induce a measureable conformational change (as recorded by DT IM-MS) in the protein, at the concentration and timescale necessary for gas-phase ion formation. This may be mainly due to the more direct method of

ion mobility separations employed in the present work, using measurements of linear drift-tube technology, as opposed to travelling-wave separations that lead to excessive heating of ions[60].

Events described in this chapter are highly complex and many-fold, and require further insights at wider range of conditions used. The main findings are summarised below.

❖ A qualitative agreement has been confirmed for earlier observations that the number of HI adducted to proteins correlate well with the number of available basic sites on it. The highest adduction levels are observed for cytochrome *c* (24 available basic residues), followed by lysozyme (19 basic residues) and BPTI (10 basic residues).

❖ HI adduction is highly charge-state dependent for lysozyme, and the opposite is true for cytochrome *c*; BPTI exhibits intermediary behaviour.

❖ Buffer gas temperature has a big effect on conformations and salt adduction in case of lysozyme and cytochrome *c*; for BPTI, lowering of the temperature has a more noticeable effect than increasing it.

❖ Increased temperature (in reference to 300 K) leads to increased CCS of lysozyme and decreased CCS of cytochrome *c* and BPTI.

❖ For lysozyme and cytochrome *c*, more HI are retained at lower temperature; for BPTI, HI are retained on higher charge states better at higher temperature. Additions of salt ions to lysozyme 'counter-act' Coulombic repulsion effect; whereas for cytochrome *c*, HI adducts have a

very limited effect on conformations; singly-adducted BPTI is a preferred arrangement across all temperatures.

❖ Lowered temperature (in reference to 300 K) improves ion mobility resolution for lysozyme only. Cross sections of lysozyme decrease with each HI added (with a few anomalies). The higher the charge state of cytochrome *c* ion, the fewer HI molecules are required to cause conformational distribution compaction. The higher the charge state of BPTI ion, the more it responds to addition of the first HI; the lower the charge state of BPTI ion, the more it responds to addition of the second HI (a decrease in CCS is observed in both cases).

❖ This is the first report of 'cold denaturation' for proteins in the gas phase, suggesting that this effect is intrinsic to the protein fold.

5.6 References

1. Kunz, W., Henle, J. & Ninham, B. W. (2004). 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. *Curr. Opin. Colloid Interface Sci.* **9**, 19-37.

2. Bauduin, P., Nohmie, F., Touraud, D., Neueder, R., Kunz, W. & Ninham, B. W. (2006). Hofmeister specific-ion effects on enzyme activity and buffer pH: Horseradish peroxidase in citrate buffer. *J. Mol. Liq.* **123**, 14-19.

3. Broering, J. M. & Bommarius, A. S. (2005). Evaluation of Hofmeister Effects on the Kinetic Stability of Proteins. *J. Phys. Chem. B* **109**, 20612-20619.

4. Curtis, R. A. & Lue, L. (2006). A molecular approach to bioseparations: Protein–protein and protein–salt interactions. *Chem. Eng. Sci.* **61**, 907-923.

5. Collins, K. D. (2004). Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Methods* **34**, 300-311.

6. Geisler, M., Pirzer, T., Ackerschott, C., Lud, S., Garrido, J., Scheibel, T. & Hugel, T. (2008). Hydrophobic and Hofmeister effects on the adhesion of spider silk proteins onto solid substrates: an AFM-based single-molecule study. *Langmuir* **24**, 1350-5.

7. Zhang, Y. & Cremer, P. S. (2009). The inverse and direct Hofmeister series for lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15249-53.

8. Zhang, Y. & Cremer, P. S. (2006). Interactions between macromolecules and ions: the Hofmeister series. *Curr. Opin. Chem. Biol.* **10**, 658-663.

9. Tanaka, K., Ido, Y., Akita, S., Yoshida, Y. & Yoshida, T. (1987). *Proc. Second Japan-China Symp. Mass Spectrom.*.

10. Karas, M. & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**, 2299-2301.

11. Meng, C. K., Mann, M. & Fenn, J. B. (1988). Of protons or proteins. *Z. Phys. D: At. Mol. Clusters* **10**, 361-368.

12. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989). Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **246**, 64-71.

13. Mirza, U. A., Cohen, S. L. & Chait, B. T. (1993). Heat-induced conformational changes in proteins studied by electrospray ionization mass spectrometry. *Anal. Chem.* **65**, 1-6.

14. Light-Wahl, K. J., Schwartz, B. L. & Smith, R. D. (1994). Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **116**, 5271-8.

15. Chowdhury, S. K., Katta, V. & Chait, B. T. (1990). Probing conformational changes in proteins by mass spectrometry. *J. Am. Chem. Soc.* **112**, 9012-13.

16. Loo, J. A. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16**, 1-23.

17. Kebarle, P. (2000). A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.* **35**, 804-817.

18. Kebarle, P. & Verkerk, U. H. (2009). Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898-917.

19. Heck, A. J. R. & van den Heuvel, R. H. H. (2004). Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368-389.

20. Johnson, R. S., Martin, S. A., Biemann, K., Stults, J. T. & Watson, J. T. (1987). Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* **59**, 2621-2625.

21. Rostom, A. A. & Robinson, C. V. (1999). Detection of the Intact GroEL Chaperonin Assembly by Mass Spectrometry. *J. Am. Chem. Soc.* **121**, 4718-4719.

22. Lorenzen, K., Versluis, C., van Duijn, E., van den Heuvel, R. H. H. & Heck, A. J. R. (2007). Optimizing macromolecular tandem mass spectrometry of large non-covalent complexes using heavy collision gases. *Int. J. Mass Spectrom.* **268**, 198-206.

23. Rodgers, M. T. & Armentrout, P. B. (2000). Noncovalent metal-ligand bond energies as studied by threshold collision-induced dissociation. *Mass Spectrom. Rev.* **19**, 215-247.

24. Despa, F., Orgill, D. P. & Lee, R. C. (2005). Molecular crowding effects on protein stability. *Ann. N. Y. Acad. Sci.* **1066**, 54-66.

25. Merenbloom, S., Flick, T., Daly, M. & Williams, E. (2011). Effects of Select Anions from the Hofmeister Series on the Gas-Phase Conformations of Protein Ions Measured with Traveling-Wave Ion Mobility Spectrometry/Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **22**, 1978-1990.

26. Han, L., Hyung, S.-J., Mayers, J. J. S. & Ruotolo, B. T. (2011). Bound Anions Differentially Stabilize Multiprotein Complexes in the Absence of Bulk Solvent. *J. Am. Chem. Soc.* **133**, 11358-11367.

27. Kruger, R. & Karas, M. (2002). Formation and fate of ion pairs during MALDI analysis: anion adduct generation as an indicative tool to determine ionization processes. *J. Am. Soc. Mass Spectrom.* **13**, 1218-1226.

28. http://www.rcsb.org/pdb/home/home.do. RCSB Protein Data Bank.

29. Whitney, S. E., Wilkinson, S. & Muller, R. (2012). Compositions for stabilizing DNA, RNA and proteins in blood and other biological samples during shipping and storage at ambient temperatures, pp. 203pp. Biomatrica, Inc., USA .

30. Lund, M. & Jungwirth, P. (2008). Patchy proteins, anions and the Hofmeister series. *J. Phys. Condens. Matter* **20**, 494218.

31. Zhou, P., Tian, F., Zou, J., Ren, Y., Liu, X. & Shang, Z. (2010). Do Halide Motifs Stabilize Protein Architecture? *J. Phys. Chem. B* **114**, 15673-15686.

32. Lund, M., Vacha, R. & Jungwirth, P. (2008). Specific ion binding to macromolecules: Effects of hydrophobicity and ion pairing. *Langmuir* **24**, 3387-3391.

33. Gokarn, Y. R., Fesinmeyer, R. M., Saluja, A., Razinkov, V., Chase, S. F., Laue, T. M. & Brems, D. N. (2011). Effective charge measurements reveal selective and preferential accumulation of anions, but not cations, at the protein surface in dilute salt solutions. *Protein Sci.* **20**, 580-587.

34. Flick, T., Merenbloom, S. & Williams, E. (2011). Anion Effects on Sodium Ion and Acid Molecule Adduction to Protein Ions in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **22**, 1968-1977.

35. Verkerk, U. H. & Kebarle, P. (2005). Ion-Ion and Ion-Molecule Reactions at the Surface of Proteins Produced by Nanospray. Information on the Number of Acidic Residues and Control of the Number of Ionized Acidic and Basic Residues. *J. Am. Soc. Mass. Spectrom.* **16**, 1325-1341.

36. Breuker, K., Jin, M., Han, X., Jiang, H. & McLafferty, F. W. (2008). Top-Down Identification and Characterization of Biomolecules by Mass Spectrometry. J. Am. Soc. Mass. Spectrom. **19**, 1045-1053.

37. Breuker, K., Brüschweiler, S. & Tollinger, M. (2011). Inside Cover: Electrostatic Stabilization of a Native Protein Structure in the Gas Phase (Angew. Chem. Int. Ed. 4/2011). *Angew. Chem. Int. Ed.* **50**, 780-780.

38. Watt, S. J., Sheil, M. M., Beck, J. L., Prosselkov, P., Otting, G. & Dixon, N. E. (2007). Effect of Protein Stabilization on Charge State Distribution in Positive- and Negative-Ion Electrospray Ionization Mass Spectra. *J. Am. Soc. Mass Spectrom.* **18**, 1605-1611.

39. Mirza, U. A. & Chait, B. T. (1997). Do proteins denature during droplet evolution in electrospray ionization? *Int. J. Mass Spectrom. Ion Processes* **162**, 173-181.

40. Taborsky, G. (1979). Protein alterations at low temperatures: an overview. *Adv. Chem. Ser.* **180**, 1-26.

41. Somero, G. N. (1995). Proteins and temperature. *Annu. Rev. Physiol.* **57**, 43-68.

42. Clemmer, D. E. & Jarrold, M. F. (1997). Ion Mobility Measurements and their Applications to Clusters and Biomolecules. *J. Mass Spectrom.* **32**, 577-592.

43. Gross, D. S., Schnier, P. D., Rodriguez-Cruz, S. E., Fagerquist, C. K. & Williams, E. R. (1996). Conformations and folding of lysozyme ions in vacuo. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 3143-8.

44. Valentine, S. J., Anderson, J. G., Ellington, A. D. & Clemmer, D. E. (1997). Disulfide-Intact and -Reduced Lysozyme in the Gas Phase:

Conformations and Pathways of Folding and Unfolding. *J. Phys. Chem. B* **101**, 3891-3900.

45. Sterling, H. J., Prell, J. S., Cassou, C. A. & Williams, E. R. (2011). Protein conformation and supercharging with DMSO from aqueous solution. *J. Am. Soc. Mass Spectrom.* **22**, 1178-1186.

46. Ries-Kautt, M. M. & Ducruix, A. F. (1989). Relative effectiveness of various ions on the solubility and crystal growth of lysozyme. *J. Biol. Chem.* **264**, 745-8.

47. Siems, W. F., Wu, C., Tarver, E. E., Hill, H. H., Jr., Larsen, P. R. & McMinn, D. G. (1994). Measuring the Resolving Power of Ion Mobility Spectrometers. *Anal. Chem.* **66**, 4195-201.

48. May, J. & Russell, D. (2011). A Mass-Selective Variable-Temperature Drift Tube Ion Mobility-Mass Spectrometer for Temperature Dependent Ion Mobility Studies. *J. Am. Soc. Mass. Spectrom.* **22**, 1134-1145.

49. Dias, C. L., Ala-Nissila, T., Karttunen, M., Vattulainen, I. & Grant, M. (2008). Microscopic Mechanism for Cold Denaturation. *Phys. Rev. Lett.* **100**, 118101.

50. Akdogan, Y. & Hinderberger, D. (2011). Solvent-Induced Protein Refolding at Low Temperatures. *J. Phys. Chem. B* **115**, 15422-15429.

51. Yoshidome, T. & Kinoshita, M. (2009). Hydrophobicity at low temperatures and cold denaturation of a protein. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **79**, 030905/1-030905/4.

52. Nash, D. P. & Jonas, J. (1997). Structure of pressure-assisted cold denatured lysozyme and comparison with lysozyme folding intermediates. *Biochemistry* **36**, 14375-14383.

53. McCullough, B. J., Kalapothakis, J., Eastwood, H., Kemper, P., MacMillan, D., Taylor, K., Dorin, J. & Barran, P. E. (2008). Development of an ion mobility quadrupole time of flight mass spectrometer. *Anal. Chem.* **80**, 6336-6344.

54. Kuroda, Y., Kidokoro, S.-i. & Wada, A. (1992). Thermodynamic characterization of cytochrome c at low pH: Observation of the molten globule state and of the cold denaturation process. *J. Mol. Biol.* **223**, 1139-1153.

55. Mao, Y., Woenckhaus, J., Kolafa, J., Ratner, M. A. & Jarrold, M. F. (1999). Thermal Unfolding of Unsolvated Cytochrome c: Experiment and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **121**, 2712-2721.

56. Shelimov, K. B. & Jarrold, M. F. (1996). "Denaturation" and Refolding of Cytochrome c in Vacuo. *J. Am. Chem. Soc.* **118**, 10313-10314.

57. Faull, P. A., Korkeila, K. E., Kalapothakis, J. M., Gray, A., McCullough, B. J. & Barran, P. E. (2009). Gas-phase metalloprotein complexes interrogated by ion mobility-mass spectrometry. *Int. J. Mass Spectrom.* **283**, 140-148.

58. Davidovic, M., Mattea, C., Qvist, J. & Halle, B. (2008). Protein Cold Denaturation as Seen From the Solvent. *J. Am. Chem. Soc.* **131**, 1025-1036.

59. Shelimov, K. B., Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. (1997). Protein Structure in Vacuo: Gas-Phase Conformations of BPTI and Cytochrome c. *J. Am. Chem. Soc.* **119**, 2240-2248.

60. Merenbloom, S., Flick, T. & Williams, E. (2011). How Hot are Your Ions in TWAVE Ion Mobility Spectrometry? *J. Am. Soc. Mass. Spectrom.*, 1-10.

Chapter 6

Conclusions and outlook

6.1 Conclusions

In the presented work, mass spectrometry and ion mobility mass spectrometry were used to probe protein-ion interactions. The proteins considered here range in mass from just under 3 kDa to over 14 kDa. Mass spectrometric measurements were carried out on a commercial ToF instrument, and IM-MS, along with some MS-only measurements, was performed on in-house modified instrument (MoQ-ToF). Both were equipped with a nESI ion source for gentle transfer of non-covalent complexes into the gas phase. For all these systems, attachment of a single ion effected a pronounced conformational change: for the peptides (vCP1 and ZiCop) – conformational rearrangement or switching between two distinctly different folds was observed, and for the proteins (lysozyme, cytochrome *c* and BPTI) – stabilisation of the fold occurred as a function of charge state and temperature. The principal discoveries of this thesis are summarised below.

In ***Chapter 3***, a consensus zinc finger peptide vCP1 was characterised. It was demonstrated that the gas-phase platform (MS and IM-MS) offers an attractive competitive route to solution-phase investigations. Binding selectivity and specificity of vCP1 to Zn^{2+} and Co^{2+} were investigated and compared to Ca^{2+} and Cu^{2+}. It was established that vCP1 preferentially binds Zn^{2+} and Co^{2+} and acquires conformational stability as a result. Collision cross sections of vCP1 were measured with and without metal ions bound, and the values were found to be very similar. Experimental CCS compared favourably with those obtained from MD simulations, giving confidence in both datasets. It was revealed that geometries of apo- and zinc-bound peptide were similarly compact, with the major conformational differences observed in the spatial arrangements of the metal-chelating residues and slight loss of the α-helical region for the apo-form.

In ***Chapter 4***, a peptide-based dual switching system ZiCop was interrogated. The dual switch is designed to be responsive to zinc binding: when bound to a Zn^{2+} ion, it is expected to adopt a zinc finger conformation, whilst in an unbound state it is designed to form a coiled coil with a partner peptide. Some aspects of the gas-phase characterisation method from the preceding chapter were applied here. Similarly to findings for the vCP1, trends in metal selectivity and specificity have been established to be preferential for Zn^{2+} and Co^{2+} compared to Ca^{2+}. Competitive Zn^{2+}/Co^{2+} binding experiments revealed that the presence of Co^{2+} enhances the strength of Zn^{2+} binding by at least one order of magnitude. Analysis of mass spectra of equimolar amounts of the ZiCop and its binding partner evidenced that, in addition to the targeted hetero-dimer formation, a competitive assembly of two homo-dimers takes place. A CID investigation of the stability of coiled coil non-covalent complexes demonstrated that hetero-dimer is more stable

compared to the two homo-dimers. Collision cross section measurements confirmed increased structural stability of the zinc-bound form of ZiCop over its free form. The hetero-dimer CCS is larger than that of apo-ZiCop suggesting an elongated spatial arrangement, a finding that is corroborated by MD simulations.

In ***Chapter 5***, the conformational stability of three proteins – lysozyme, cytochrome *c* and BPTI, was probed as a function of temperature and HI adductation. Primary metrics used in this study were the relative intensities of MS signals and the CCS of proteins and protein-salt clusters. The findings suggest that the presence of the salt in the same solution as the protein is sufficient to induce a significant conformational change in some proteins studied here, at the concentration and timescale necessary for gas-phase ion formation. For lysozyme, effect of 'counter-acting' Coulombic repulsion by additions of salt ions was observed. Cytochrome *c* yielded a very limited response to salt ions additions, and the preferred stoichiometry for BPTI was a singly-adducted state. In all cases, decrease in CCS distribution (*i.e.* conformational compaction) is observed with each HI added. Temperature studies have shown that proteins respond by conformational changes to a decrease in temperature more readily than to its increase. Cold denaturation effect was observed for the first time in the gas phase.

6.2 Outlook

Offered below are some perspectives on the opportunities and challenges in the characterisation of biological assemblies in the gas phase. Mass spectrometry, enhanced by capabilities of ion mobility spectrometry, is a powerful tool to study protein structure that offers unparalleled sensitivity, specificity and speed. Insights into the tertiary structure of proteins in a solvent-free environment can be obtained by measuring collision cross sections. An additional dimension of structural arrangement of proteins and their complexes can be obtained by molecular dynamics simulations, whereby candidate high-resolution geometries can be compared and assigned to intrinsically low-resolution ion mobility data. Non-covalent complexes can be studied to a high degree of detail, yielding information on stoichiometry, binding preferences, as well as the measure of the strength of association of constituents of such complexes. While in the gas phase, complexes can be investigated at different temperatures, and information on their stability can be deduced from changes in the resulting collision cross sections. IM-MS is a potent probe for structural studies that is capable of registering single-ion events taking place in large biopolymer systems. This technique already plays an instrumental role in various disciplines such as proteomics, metabolomics and fundamental biological research. At this point, biophysical mass spectrometrists with interest in technique development are gradually shifting their efforts from demonstrating the applicability of the method to making it more available to a wider pool of potential users[1]. This shift can only be accomplished if a number of technical obstacles limiting the use of the technique are addressed, a subset of which is outlined below.

It is often difficult to retain functional conformations and complexes of proteins into the gas phase for IM-MS analysis. The challenge to preserve fragile structures from solution into the gas phase is associated with the formation gas-phase ions. Nano-electrospray emitters are extremely appealing due to their ability to gently transfer labile biological molecules and large complexes into the gas phase preserving their native state. However this comes with a lack of robustness of the method, as electrospray performance tends to deteriorate over time, mainly due to clogging, which requires constant monitoring of the process. The distribution in the spray tip geometry from batch to batch is a major issue too, resulting in diminished reproducibility of results. The overall trend is the move towards arrays of emitters that are microfabricated to high specifications, ensuring higher spray stability and enabling unattended (to a degree) operation[2].

Another bottleneck for successful application of IM-MS in biological research exists due to the method's sensitivity to instrument settings. Caution needs to be exercised when optimising parameters of mobility separations. This is especially the case for the only (so far) commercial instrument, whose performance has been optimised for sensitivity, resolution and high-throughput, often at a price of over-heating ions immediately prior and during mobility separation, causing the loss of proteins' functional form[3]. Alternative mobility separation technologies (such as linear drift tube) are needed in the market to enable expansion of applicability of IM-MS into wider areas of research. As of today, mass spectrometers containing a linear drift tube are only available to research groups who are able and willing to build their own instruments. A more proactive input form global instrument companies will be needed to ensure that a wider choice of competitive and complementary technologies becomes available on the market.

6.3 References

1. van Duijn, E. (2010). Current Limitations in Native Mass Spectrometry Based Structural Biology. *J. Am. Soc. Mass. Spectrom.* **21**, 971-978.

2. Gibson, G. T. T., Mugo, S. M. & Oleschuk, R. D. (2009). Nanoelectrospray emitters: trends and perspective. *Mass Spectrom. Rev.* **28**, 918-936.

3. Merenbloom, S., Flick, T. & Williams, E. (2011). How Hot are Your Ions in TWAVE Ion Mobility Spectrometry? *J. Am. Soc. Mass. Spectrom.*, 1-10.

Appendices

Selected properties of amino acids

Figure A 1 lists the 20 amino acids that naturally occur in proteins. Here, amino acids are arranged in rows: the polar amino acids are in the first two rows, the hydrophobic ones are in the middle two, and the amino acids classed as 'special' are in the bottom row. This is of course only one way of grouping them according to their properties. Alternatively, they can be classed as aromatic, aliphatic, positively or negatively charged, and finally by size of the side chain.

Table A 1 contains acid dissociation constants pK_a for the ionisable groups of the 20 amino acids – amino, carboxylic and, where applicable, side chain. Isoelectric points pI at 25°C are also listed.

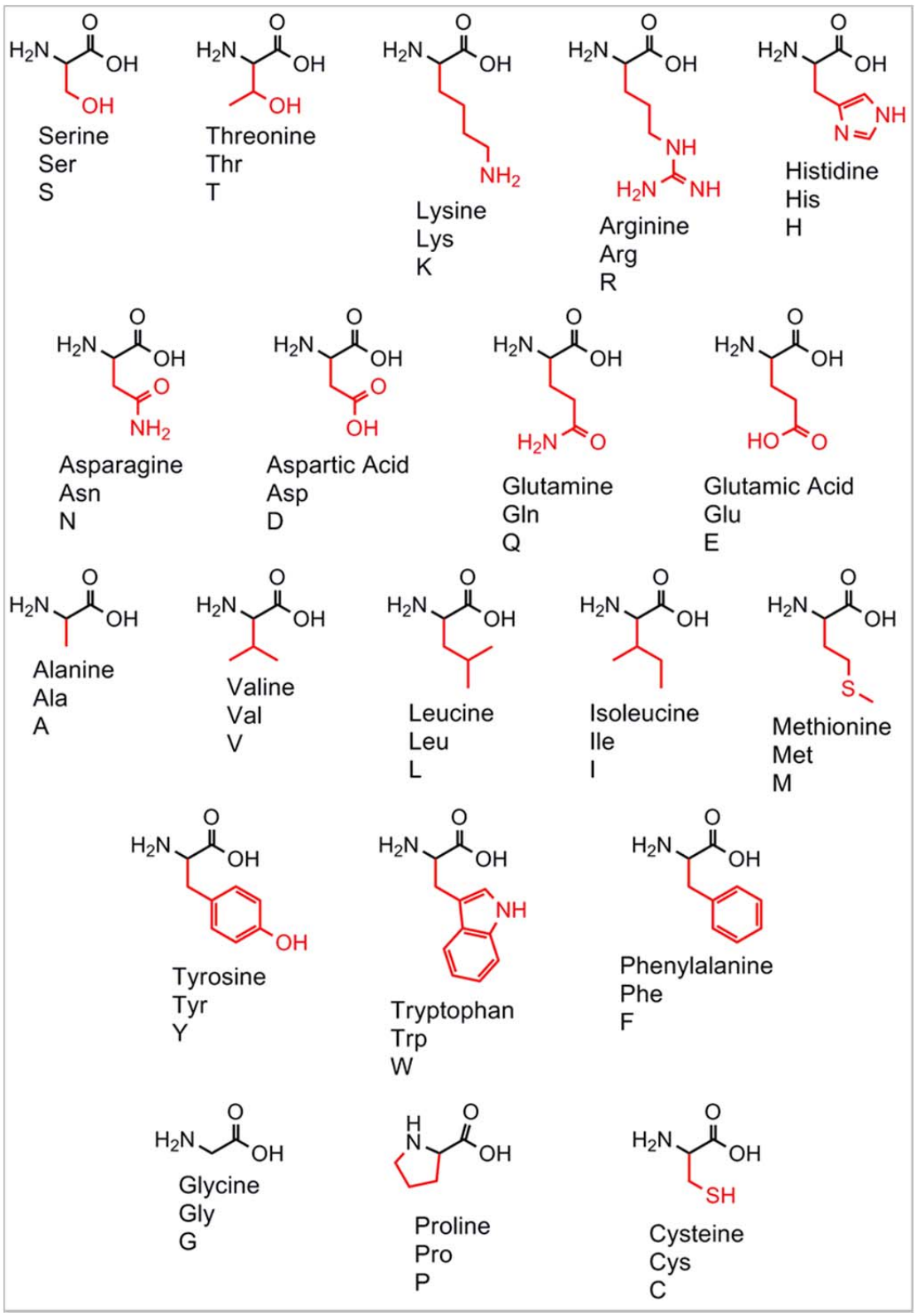


Figure A 1 *The structures of twenty common amino acids occurring in natural proteins, along with their names, three- and one-letter codes; the side chains are highlighted in red. The first two rows show the polar amino acids, the middle two the hydrophobic ones, and the bottom row shows the amino acids classed as 'special'.*

Amino acid name	Molecular mass, Da		pKa			pI at 25°C
	mono-isotopic	average	α-CO_2H	α-NH_2	R-group	
Alanine	71.04	71.08	2.35	9.87		6.11
Arginine	156.10	156.19	2.18	9.09	13.20	10.76
Asparagine	114.04	114.10	2.02	8.80		5.41
Aspartic acid	115.03	115.09	1.88	9.60	3.65	2.98
Cysteine	103.01	103.14	1.71	10.78	8.33	5.02
Glutamic acid	129.04	129.11	2.19	9.67	4.25	3.08
Glutamine	128.07	128.13	2.17	9.13		5.65
Glycine	57.03	57.05	2.34	9.60		6.06
Histidine	137.07	137.14	1.78	8.97	5.97	7.64
Isoleucine	113.09	113.16	2.32	9.76		6.04
Leucine	113.09	113.16	2.36	9.60		6.04
Lysine	128.09	128.17	2.20	8.90	10.28	9.47
Methionine	131.05	131.20	2.28	9.21		5.74
Phenylalanine	147.08	147.17	2.58	9.24		5.91
Proline	97.06	97.12	1.99	10.60		6.30
Serine	87.04	87.08	2.21	9.15		5.68
Threonine	101.05	101.10	2.15	9.12		5.60
Tryptophan	186.09	186.21	2.38	9.39		5.88
Tyrosine	163.07	163.17	2.20	9.11	10.07	5.63
Valine	99.08	99.13	2.29	9.74		6.02

Table A 1 *pK_a and pI values of amino acids. Adapted from http://www.anaspec.com.*

Selected properties of metal ions considered in this work: Zn, Co, Ca and Cu

Table A 2 summarises selected properties of divalent metal ions relevant to the study of metal coordination in synthetic peptides vCP1 and ZiCop, described in ***Chapters 3*** and ***4*** respectively. The metal central to these studies is the transition metal zinc; two other transition metals have been investigated for comparison – cobalt and copper, along with one alkaline

earth metal – calcium. Electron shell properties determine whether or not a metal binding can be monitored spectroscopically (see main text in section ***1.1.4***); ionic radius potentially has a steric effect on coordination sphere of the peptides; metal hardness governs the binding preferences (*i.e.* O-, N- or S-containing amino acid side chains).

Metal ion	**Molecular mass, Da**		**Electron shell structure (max. electrons permitted per shell: $s^2p^6d^{10}f^{14}g^{18}$)**	**Ionic radius[1], pm**	**Hardness[2]**
	mono-isotopic	**average**			
Zn^{2+}	63.93	65.39	2.8.18.2	74	borderline
Co^{2+}	58.93	58.93	2.8.15.2	54.5	borderline
Ca^{2+}	39.96	40.08	2.8.8.2	100	hard
Cu^{2+}/Cu^{1+}	62.93	63.55	2.8.18.1	73/77	borderline/soft

Table A 2 *Summary of selected properties of divalent ions used to study metal coordination in synthetic peptides vCP1 and ZiCop.*

The hardness scale separates metal ions into class A (hard), class B (soft) and borderline (intermediate)[2]. Within the borderline metal hardness range listed in the table, the following rank is observed (from hardest to softest): Zn^{2+} – Co^{2+} – Cu^{2+}. The ranking of class A metal ions' preferences for donor groups containing following atoms is as follows: O > N > S, whereas class B ranking is reversed. Thus in proteins, among functional groups sought by hard ions is carboxylate (e.g. of Asp and Glu), and by soft ions – sulphydryl (e.g. thiolate of Cys) and heterocyclic nitrogen (e.g. imidazole of His).

Metal ions and vCP1 – quantifying metal ion affinity by spectroscopic measurements

To confirm the secondary structure and metal binding of vCP1 in solution, circular dichroism (CD) spectroscopy and ultraviolet (UV) spectroscopy were used respectively. The CD spectra in the absence of metal ions confirm that the peptide was largely unstructured, whereas upon the addition of zinc the peptide folded to give a spectrum consistent with that reported for naturally occurring zinc fingers, ***Figure A 2A***. Addition of cobalt effected a similar change to that seen with zinc; in contrast, addition of copper and calcium changed the conformation of the peptide in a less pronounced manner (***Figure A 3***). The addition of cobalt to vCP1 was also followed by UV spectroscopy, ***Figure A 2B***. The spectra are consistent with Cys_2His_2 tetrahedral coordination of the metal[3], and a binding constant of 0.8 μM (***Figure A 4***). A competitive binding assay for zinc against the cobalt : vCP1 complex revealed sub-nM binding of zinc. These data confirm that vCP1 is a good model for metal binding studies under the solution conditions needed to best preserve intact complexes into the solvent-free environment of a mass spectrometer.

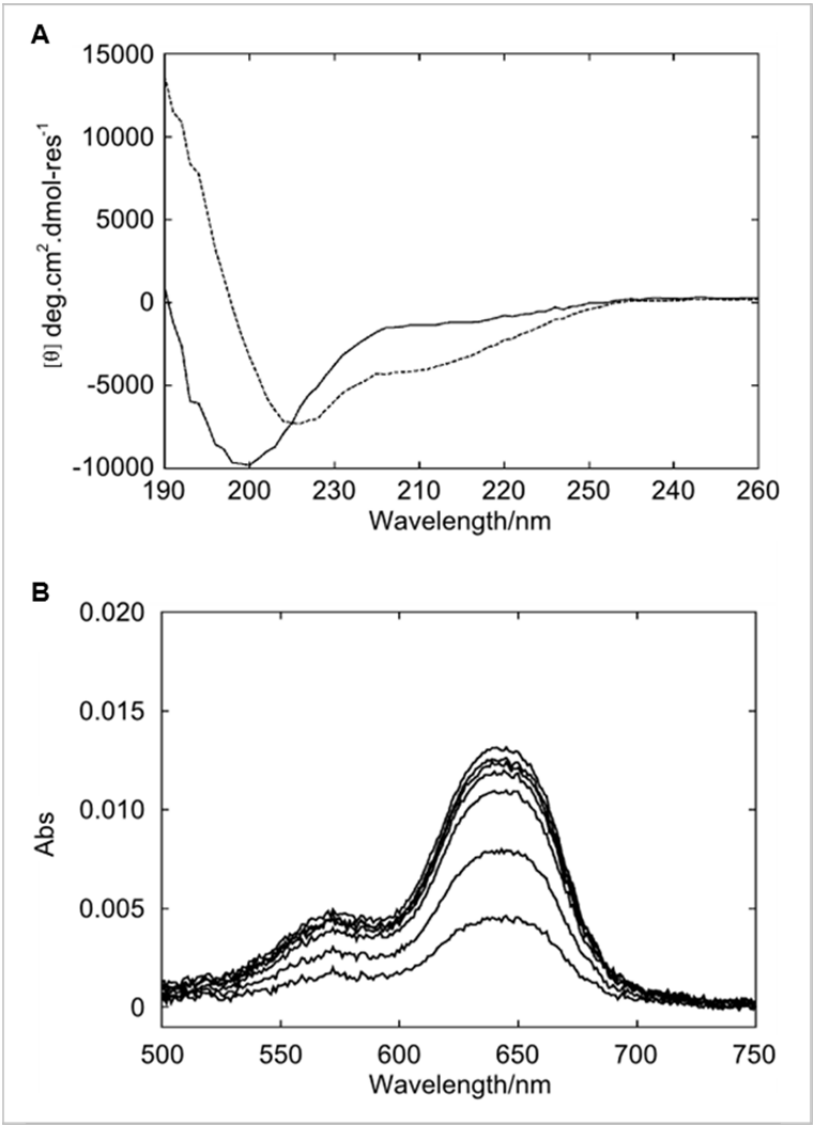


Figure A 2 *Solution-phase spectra for vCP1. (**A**) CD spectra of 25 μM peptide in the absence (solid line) and presence (dotted line) of 50 μM zinc. (**B**) UV spectra for a titration of 4.4 – 47.4 μM Co^{2+} into 11.9 μM vCP1. Conditions: 20 mM ammonium acetate; 5% isopropanol; pH 7.2; 500 μM TCEP.*

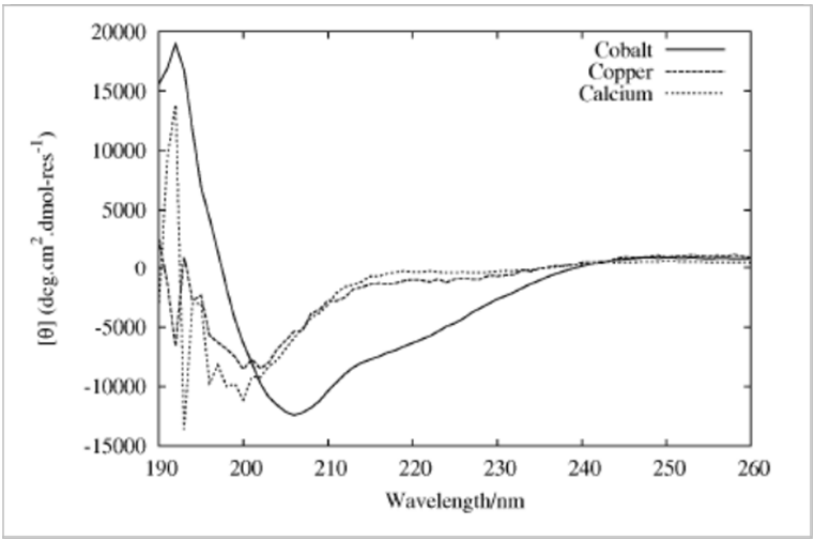


Figure A 3 *CD spectra of 25 μM vCP1 in the presence of 50 μM cobalt, copper and calcium. The cobalt spectrum is similar (although not identical) to that obtained using zinc. The copper and calcium spectra are more apo-like.*

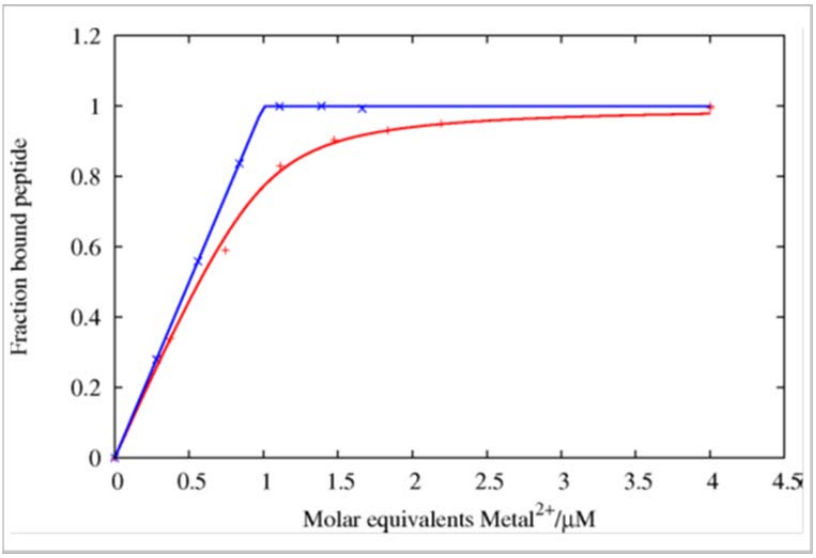


Figure A 4 *Fits to the binding data for the addition of zinc (blue) and cobalt (red) to vCP1 obtained from UV data. The K_d for cobalt is estimated to be 0.8 μM and that for zinc is tight and sub-nanomolar.*

Molecular modelling of vCP1

holo-vCP1

Figure A 5 shows the time series of the root mean square deviation (RMSD), the radius of gyration (R_g) and the average CCS calculated using trajectory method. When RMSD from the backbone heavy atoms is considered, it is evident that during the simulation the peptide conformation does not diverge much from the initial structure (the RSMD is always less than 4 Å). The R_g is calculated for all the atoms and testifies that the dimension of the peptide is roughly the same, except for a slight collapse at 8 ns. At or around this point, hydrophilic (charged) side chains form salt bridges between themselves and with charged atoms of the backbone[4]. In addition to the TM, the other two methods for calculating CCS were implemented for comparison – PA and EHSS. The trend of the calculated collision cross sections is the same for the three implemented methods; the absolute values reflect the expected deviations: PA method underestimated Ω_{avg} by 12%, EHSS and TM are both near the experimental values. The time averages of the cross sections are: 431 $Å^2$, 507 $Å^2$ and 495 $Å^2$ for PA, EHSS and TM respectively. ***Figure 3-5*** in ***Chapter 3*** shows a representative structure from this time course by trajectory method.

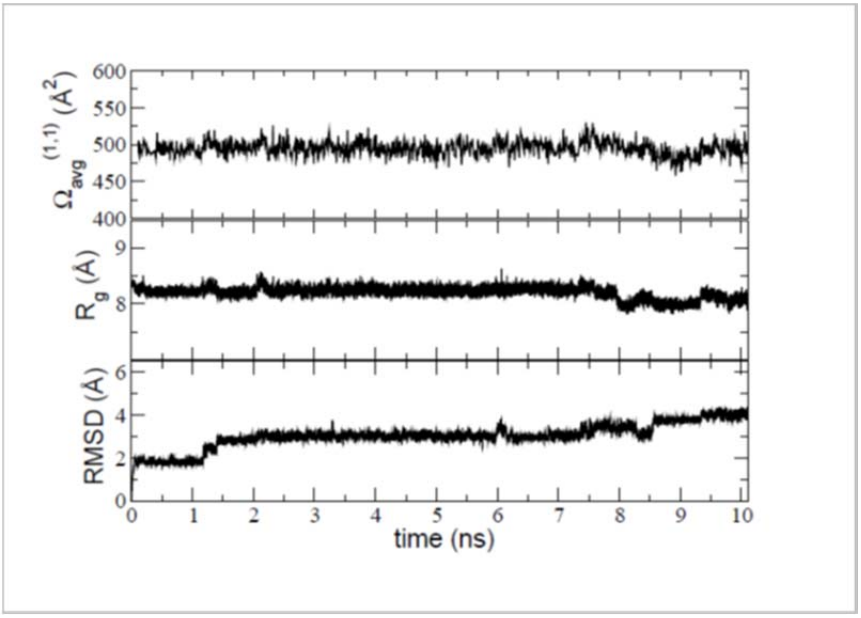


Figure A 5 *Holo-vCP1 time series of the average TM collision cross section (upper graph), radius of gyration (middle graph) and root mean square deviation (lower graph). Calculation of the 'all time' values was neglected for the heating process (first 10 ps).*

apo-vCP1

The two strategies applied to molecular modelling of the apo-vCP1 are described in section ***2.7.1*** of ***Chapter 2*** and the result of the second strategy was reported in ***Figure 3-5*** of ***Chapter 3***.

In strategy 1, a simulated annealing (SA) protocol was applied starting from a fully extended peptide. A 100 of minimised structures were obtained and for each of them the TM collision cross section was calculated (*via* MobCal). The plot of CCS values *versus* energy values is given in ***Figure A 6***. Resulting average TM collision cross for these 100 structures Boltzmann weighted by their energy equalled 519 Å^2. A time evolution MD of the CCS was performed on two representative structures as indicated in ***Figure A 6*** with red circles. First is the one with the lowest energy value and the second is

the one whose TM collision cross section value is the nearest to that experimentally measured. The time series of these two runs are given in ***Figure A 7*** and ***Figure A 8*** respectively. Their corresponding time-averaged TM collision cross sections are 512 Å^2 and 497 Å^2.

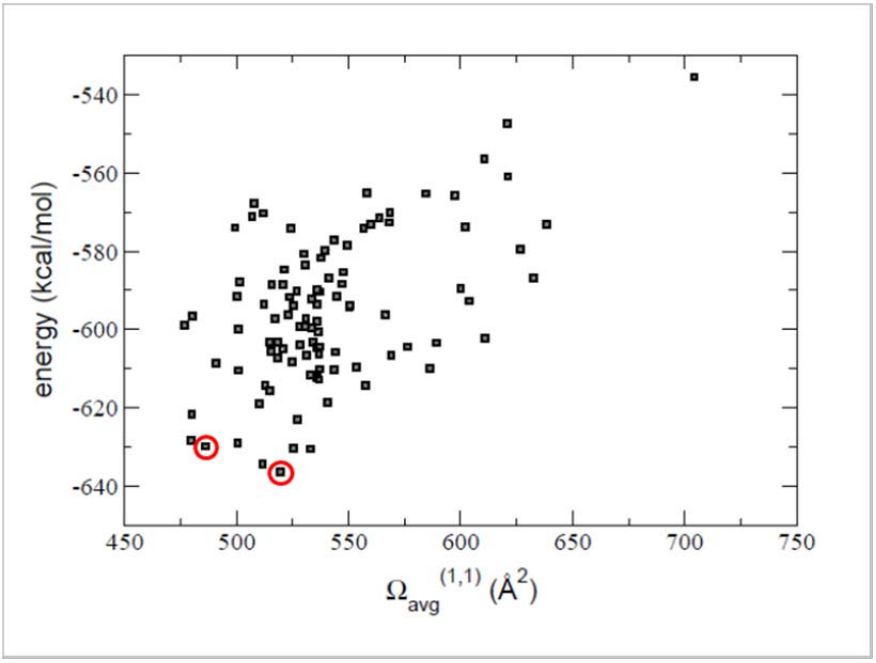


Figure A 6 *Scatter plot of the collision cross section (by TM) versus energy for the 100 simulated annealing structures. Structures that were selected for subsequent MD are marked with red circles.*

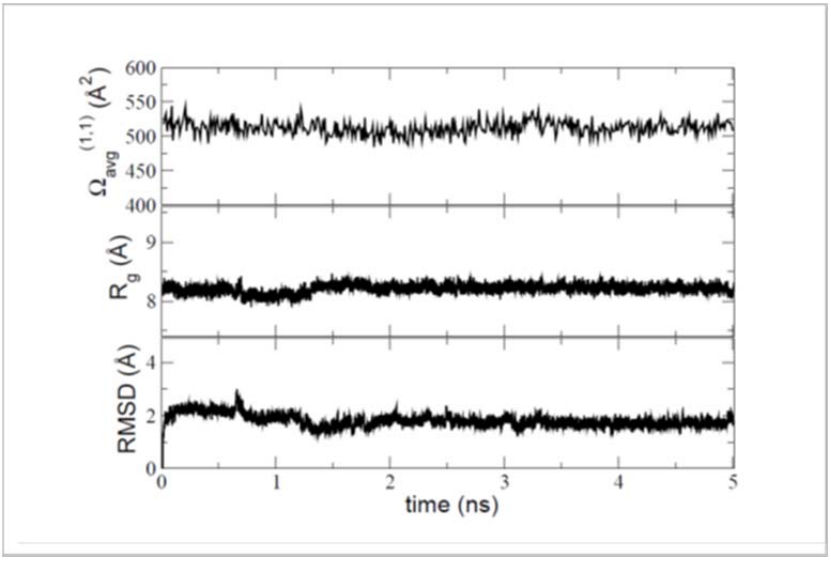


Figure A 7 *Time series for the apo-vCP1 structure with the lowest value of total energy among the 100 simulated annealing structures. Average TM CCS (upper graph), radius of gyration (middle graph) and root mean square deviation (lower graph). Calculation of the 'all time' series was neglected for the heating process (first 10 ps).*

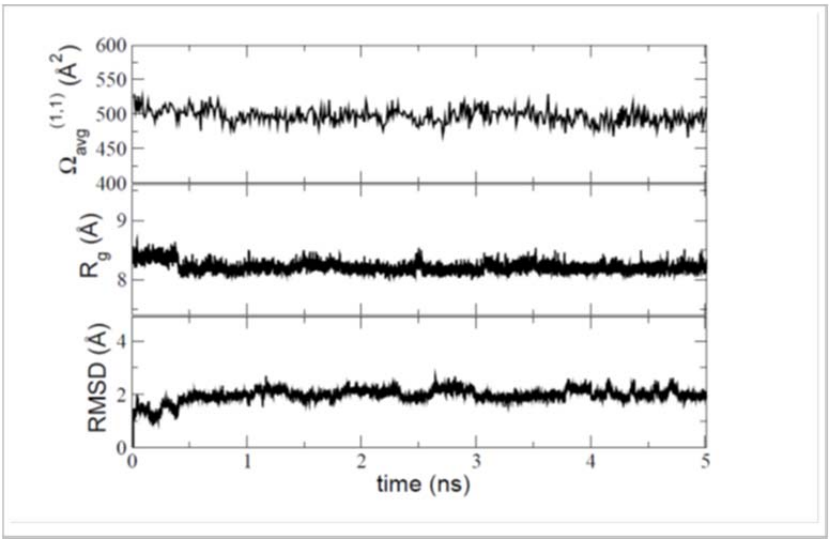


Figure A 8 *Time series for the apo-vCP1 simulated annealing structure with the CCS closest to the experimental one. Average TM CCS (upper graph), radius of gyration (middle graph) and root mean square deviation (lower graph). Calculation of the 'all time' series was neglected for the heating process (first 10 ps).*

In the second strategy the zinc atom was removed from the holo structure and an MD simulation run for to 10 ns. The strategy 2 time series are shown in ***Figure A 9*** and, in this case, the time averaged TM collision cross section is

501 Å^2. This value, along with a representative structure, is reported in ***Figure 3-5*** in ***Chapter 3*** of the main text. Indeed, when the CCS of 501 Å^2 is compared to the CCS obtained by using the first strategy (512 Å^2 and 497 Å^2), it is evident that it falls between them and can be considered to be in the same range.

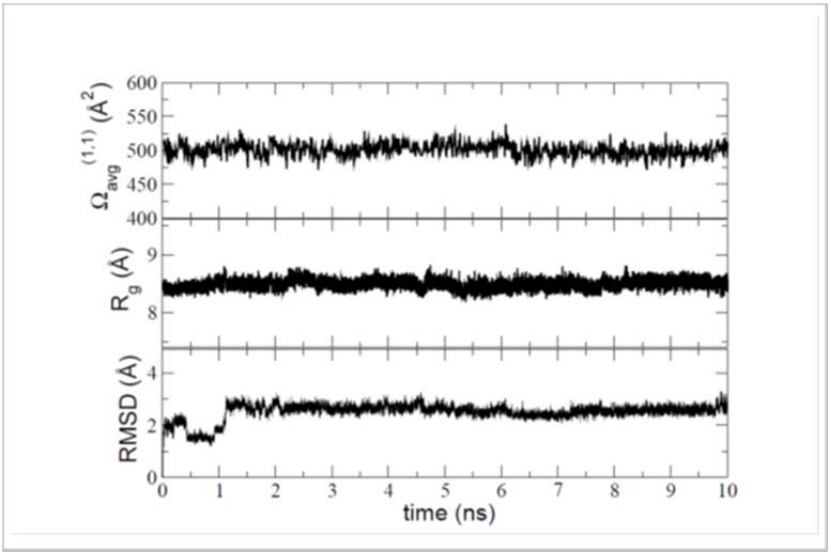


Figure A 9 *Time series for the apo-vCP1 system derived by direct removal of the zinc atom from the holo structure. Average TM CCS (upper graph), radius of gyration (middle graph) and root mean square deviation (lower graph). Calculation of the 'all time' series was neglected for the heating process (first 10 ps).*

Metal ions and ZiCop – quantifying metal ion affinity by spectroscopic measurements

Figure A 10 shows the results of using UV-visible spectroscopy to probe the metal binding properties of ZiCop. ZiCop was found to bind cobalt with a K_d of 7.5 μM, ~10-fold less tight than vCP1. This work has been carries out by the collaborating team, and below are their findings.

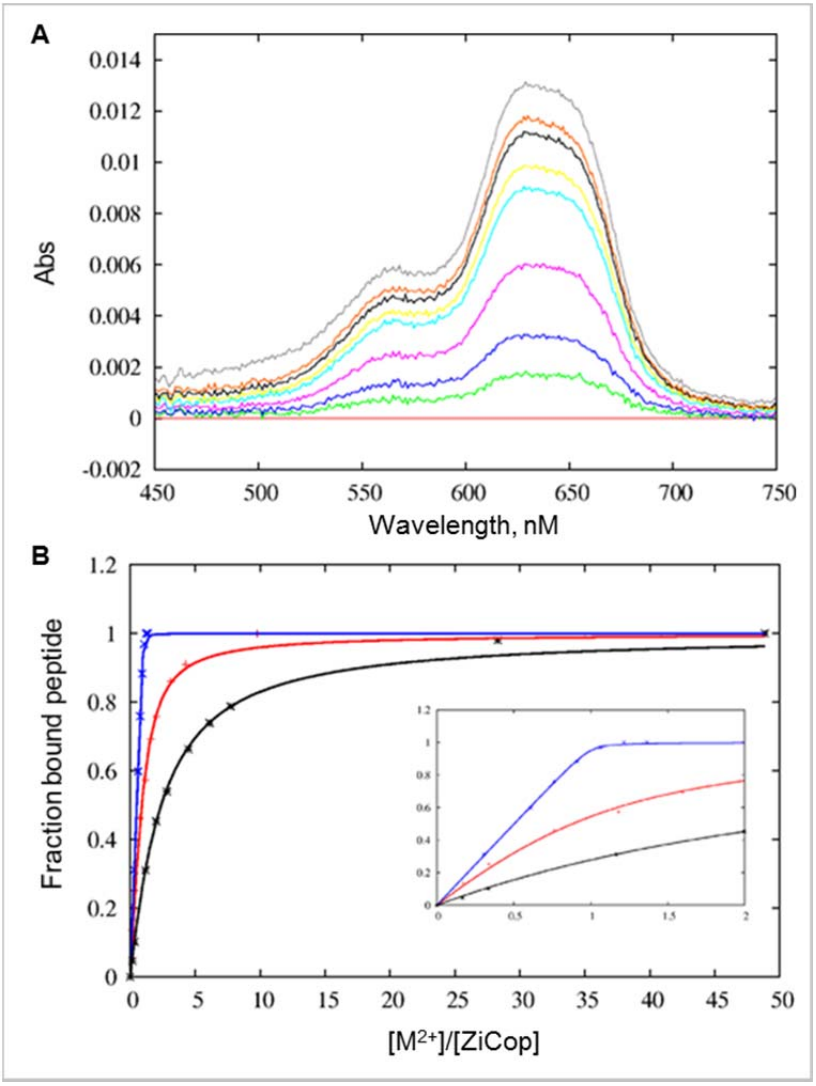


Figure A 10 *Metal-binding properties of ZiCop as monitored by UV-visible spectroscopy. (**A**) The spectra obtained when cobalt was added to ZiCop (peptide conc. 19.57 μM, cobalt concentration ranging from 1 to 190 μM). (**B**) Titrations followed at 630 nm plotted as a function of added cobalt. Red – cobalt binding to 19.6 μM ZiCop (K_d ≈ 7.5 μM), Blue – zinc displacing 278 μM cobalt from 14.5 μM ZiCop (K_d ≈ 40 nM), black – cobalt binding to ZiCop in the presence of 21.2 μM partner peptide (K_d ≈ 28 μM). The inset shows the binding properties at lower metal concentrations more clearly. All spectra recorded at pH 7.2.*

Adding zinc chloride to a solution of cobalt-bound ZiCop resulted in the displacement of cobalt from the peptide. Metal exchange rates were observed to be slow – probably due to slow cobalt off-rates – so after each addition of zinc chloride the solution was left to equilibrate for 30 minutes. The concentration of ZiCop was estimated as being 19.6 μM; zinc binding, however, appeared to be complete after the addition of around 14.5 μM zinc

chloride. Estimating peptide concentration by UV measurements is known to be prone to error[5], though the apparent 25% error observed here is particularly high. It is possible, therefore, that some 2:1 ZiCop:Zn complexes exist. When the binding curves were fit using an altered ZiCop concentration of 14.5 μM, a K_d indicative of tight binding was observed; even in the presence of an excess of cobalt the binding was too tight for an exact K_d to be assigned.

Experimental and calculated CCS values for lysozyme, cytochrome *c* and BPTI

Table A 3 summarises collision cross sections determined for all ions observed at 300 K buffer gas temperature for the three model proteins studied in ***Chapter 5*** (lysozyme, cytochrome *c* and BPTI). The values in the table indicate drift peak maxima for each observed charge state that have been plotted as a function of adduction level. ***Table A 4*** summarises crystal structure values for the three apo-proteins at 260, 300 and 360 K.

Lysozyme

CS	Peak 1				Peak 2			
	M	M+1I	M+2I	M+3I	M	M+1I	M+2I	M+3I
7		1487±214	1370±274	1339±283		1663±349	1584±262	1582±202
8	1289±16	1346±17	1322±17	1301±20	1602±155	1664±80	1581±138	1573±136
9	1492±38	1488±86	1446±88	1433±78	1771±127	1677±150	1633±123	1600±99
10	1710±140	1646±79	1525±69	1437±39	1847±92	1796±93	1761±89	1764±58
11	1798±38	1777±76	1658±97		2038±60	1997±40	1956±65	

Cytochrome *c*

CS	Peak 1				Peak 2			
	M	M+1I	M+2I	M+3I	M	M+1I	M+2I	M+3I
6	1043±45	1022±104			1188±77	1262±60		
7	1280±38	1399±129	1270±56	1351±93	1631±164	1518±136	1493±201	1659±185
8	1396±24	1306±69	1473±104	1523±56	1867±76	1799±113	1759±162	1812±175
9	1515±146	1613±207			1969±144	1930±223		

BPTI

CS	Peak 1			Peak 2		
	M	M+1I	M+2I	M	M+1I	M+2I
4	645±26	645±22	623±31	775±15	775±17	741±28
5	742±12	731±11	733±7	798±16	777±12	775±10
6	811±22	790±11		848±14	817±10	
7	920±4	872±5		946±16	890±6	

Table A 3 *Experimental numerical CCS values at 300 K for lysozyme, cytochrome c and BPTI as plotted in* ***Figure 5-8****,* ***Figure 5-12*** *and* ***Figure 5-16*** *in* ***Chapter 5****. Each value represents the average (± standard error) of three experimental repeats conducted on different days.*

Protein, charge state	Buffer gas temperature, K		
	260	300	360
Lysozyme, 8+	1471.9	1461.6	1445.7
Cytochrome C, 6+	1305.4	1291.8	1276.9
BPTI, 6+	921.2	910.8	901.8

Table A 4 *Calculated in MobCal numerical CCS values for minimised crystal structures of lysozyme, cytochrome c and BPTI as plotted in ATD and CCS diagrams in* ***Chapter 5****. The following PDB structures are used: 3AW6 (for lysozyme), 1HRC (for cytochrome c) and 1BPI (for BPTI).*

Molecular modelling of lysozyme, cytochrome *c* and BPTI

To analyse the dynamical evolution of the gas-phase conformational rearrangement, three 5-ns *in vacuo* simulations were run for each protein at the three different experimental temperatures (260 K, 300 K and 360 K). The minimised structures were used as input files, an 'infinite' radial cut-off was implemented, bonds involving hydrogen atoms were kept constrained at their equilibrium length by *shake* algorithm[6] and a time step of 1.0 fs was utilised. During dynamics the temperature was kept constant using Langevin algorithm, with a collision coefficient equal to 2.0 ps^{-1}. The resulting backbone root mean square deviation (RMSD) and radius of gyration (R_g) are displayed in ***Figure A 11***.

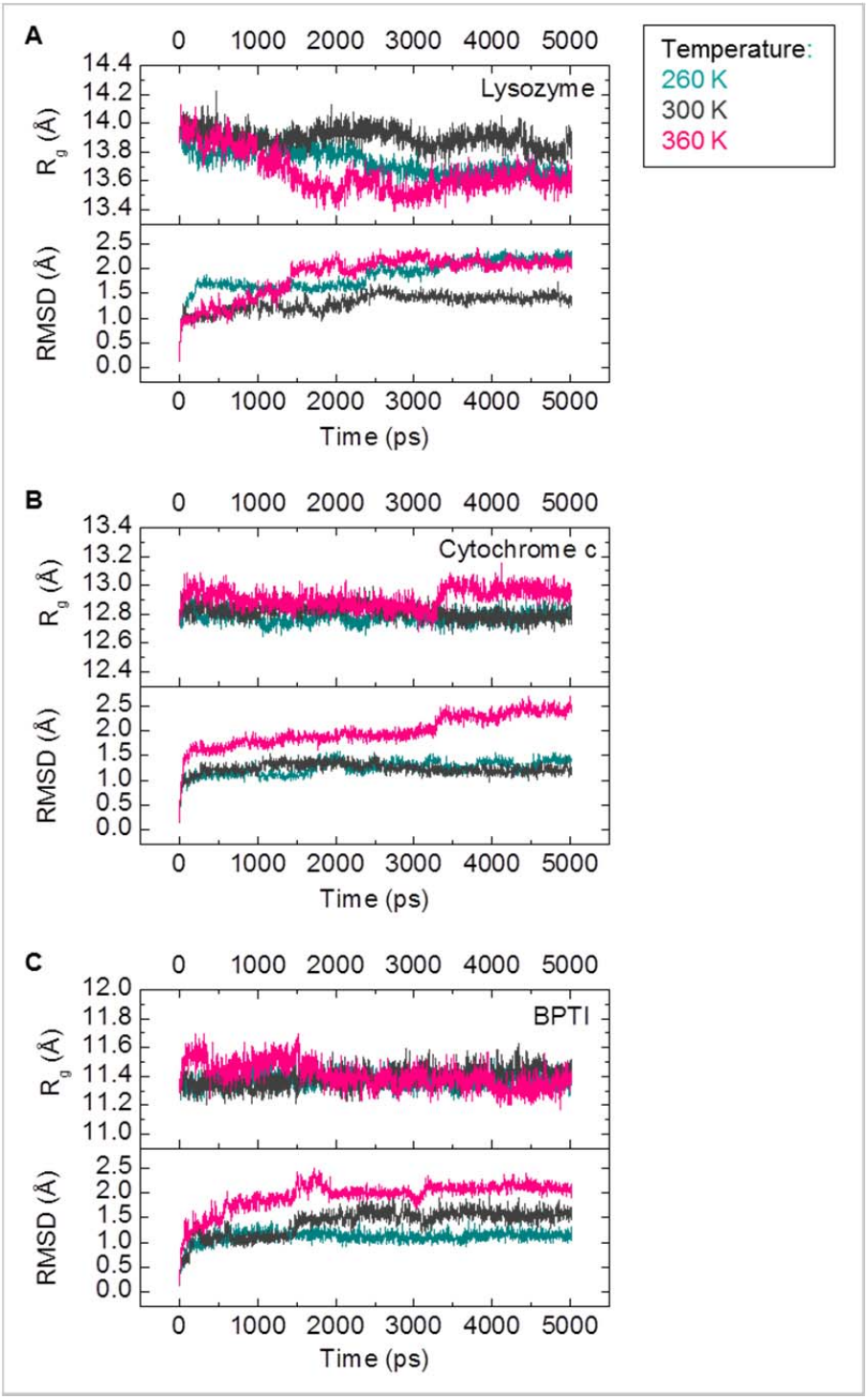


Figure A 11 *RMSD and R_g trends of the three model proteins: (**A**) lysozyme, (**B**) cytochrome c, (**C**) BPTI, at the three simulation temperatures – turquoise, grey and pink traces for T = 260 K, 300 K and 360 K respectively.*

Selected physical constants, symbols and units

The three tables below contain selected fundamental physical constants (along with their symbols and units) that were used in this work.

Symbol	Name	Value (SI units)
Da	atomic mass unit	$1.66054 \cdot 10^{-27}$ kg
e	elementary charge	$1.6021765 \cdot 10^{-19}$ C
eV	electronvolt	$96.4853 \cdot 10^{3}$ J mol^{-1}
k_B	Bolzmann constant	$1.3806503 \cdot 10^{-23}$ J K
m_e	electron mass	$9.1093897 \cdot 10^{-31}$ kg
m_p	proton mas	$1.6726231 \cdot 10^{-27}$ kg
N_A	Avogadro constant	$6.0221413 \cdot 10^{23}$ mol^{-1}
P_0	standard pressure	101,324.72 Pa (= 760 Torr)
R	molar gas constant	8.314510 J mol^{-1} K
T_0	standard temperature	273.15 K (= 0°C)
Td	Townsend	$1 \cdot 10^{-17}$ V cm^2

Table A 5 Major physical constants.

Symbol	Name	SI unit symbol
E	electric field strength	$V\ m^{-1}$
E_k	kinetic energy	$J = kg\ m^2\ s^{-2}$ ($1\ J = 2.39 \cdot 10^{-4}$ kcal $= 6.24 \cdot 10^{18}$ eV)
K	(ion) mobility	$m^2\ V^{-1}\ s^{-1}$
K_0	reduced (ion) mobility	$m^2\ V^{-1}\ s^{-1}$
K_d	dissociation constant	M
L, x	distance	m ($1\ m = 1 \cdot 10^{10}$ Å^2)
m	mass	kg
m/z	mass-to-charge ratio	
MW	molecular weight	Da
N	number density	m^{-3}
P	pressure	Pa ($1\ Pa = 10^{-5}$ bar $= 9.8692 \cdot 10^{-6}$ atm $= 7.5006 \cdot 10^{-3}$ Torr $= 145.04 \cdot 10^{-6}$ psi)
pH	activity of the (solvated) hydrogen ion	
pI	isoelectric point	
q	net charge	C
R_g	radius of gyration	m
t	time	s
T	temperature	K ($1\ K = 1\ °C$; [K] = [°C] + 273.15)
t_0	dead time	s
t_a	arrival time	s
t_d	drift time	s
U	DC potential	V
v	velocity	$m\ s^{-1}$
V	electric potential difference; peak amplitude of RF potential	V
z	nominal charge (charge number)	
μ	reduced mass	kg
ν	RF	$Hz = s^{-1}$
Φ_0	total electric potential	V
ω	angular frequency	$Hz = s^{-1}$
Ω	momentum transfer integral (collision cross section)	Å$^2 = 1 \cdot 10^{-20}\ m^2$ (Å $= 1 \cdot 10^{-10}$ m)

Table A 6 Major physical symbols and units.

Prefix	Symbol	Factor 10^n
tera	T	10^{12}
giga	G	10^{9}
mega	M	10^{6}
kilo	k	10^{3}
hecto	h	10^{2}
deca	da	10^{1}
deci	d	10^{-1}
centi	c	10^{-2}
milli	m	10^{-3}
micro	μ	10^{-6}
nano	n	10^{-9}
pico	p	10^{-12}
femto	f	10^{-15}
atto	a	10^{-18}

Table A 7 Metric system unit prefixes.

Published work

Work presented in Chapter 3 was published in Chemical Communications journal and the article is reproduced here in full by permission of The Royal Society of Chemistry. The reprint can be found at the end of this thesis.

References

1. Shannon, R. D. (1976). Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr., Sect. A* **A32**, 751-67.

2. Nieboer, E. & Richardson, D. H. S. (1980). The replacement of the nondescript term 'heavy metals' by a biologically and chemically significant classification of metal ions. *Environ. Pollut., Ser. B: Chem. Phys.* **1**, 3-26.

3. Shi, Y., Beger, R. D. & Berg, J. M. (1993). Metal binding properties of single amino acid deletion mutants of zinc finger peptides: studies using cobalt(II) as a spectroscopic probe. *Biophys. J.* **64**, 749-753.

4. Breuker, K. & McLafferty, F. W. (2008). Stepwise evolution of protein native structure with electrospray into the gas phase, 10E-12 to 10E2 s. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18145-18152.

5. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). How to Measure and Predict the Molar Absorption-Coefficient of a Protein. *Protein Sci.* **4**, 2411-2423.

6. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327-41.

Metal binding to a zinc-finger peptide: a comparison between solution and the gas phase†‡

Yana Berezovskaya,[a] Craig T. Armstrong,[b] Aimee L. Boyle,[b] Massimiliano Porrini,[a] Derek N. Woolfson*[bc] and Perdita E. Barran*[a]

Received 19th July 2010, Accepted 1st September 2010
DOI: 10.1039/c0cc02445g

Solution-phase spectroscopy and mass spectrometry are used to probe interactions between divalent metal ions and a synthetic Cys_2His_2 zinc-finger peptide (vCP1). Both methods provide the same order of binding affinity, zinc ≥ cobalt ≫ copper ≫ calcium. Collision-cross-section measurements show that both apo and holo forms are compact. This is corroborated by molecular-dynamics simulations.

Advances in mass spectrometry (MS) have placed it at the forefront of techniques for characterising proteins. For the past three decades, mass spectrometry has been increasingly used to provide masses of single proteins, and of protein fragments as part of proteomics and protein-sequencing studies. More recently, developments in soft ionization methods have allowed researchers to observe protein complexes in the gas phase, signifying a move from studying single proteins to protein systems. Moreover, it has been possible to record the time taken for proteins to move through inert gases in a drift cell using ion mobility mass spectrometry (IM-MS), allowing users to calculate collision cross sections and hence obtain valuable information on the conformations adopted in the gas phase.

We employ mass spectrometry in combination with the related technique of ion mobility mass spectrometry to elucidate conformations of charged biomolecules in the gas phase.[2,3] IM-MS has already proven useful in fundamental studies of sequence-to-structure relationships in polypeptides, since structures observed *in vacuo* are defined intrinsically by the sequence.[4,5] Ion-mobility instrumentation determines the time it takes ions to pass through a drift cell containing an inert gas under the influence of a weak electric field.[6] The mobility of an ion is the constant of proportionality between the velocity at which it moves and the applied electric field, and is inversely proportional to the collision cross section (Ω) of the gas-phase ion:

$$K_0 = \frac{3ze}{16N}\left(\frac{2\pi}{\mu k_B T}\right)^{1/2}\frac{1}{\Omega} \qquad (1)$$

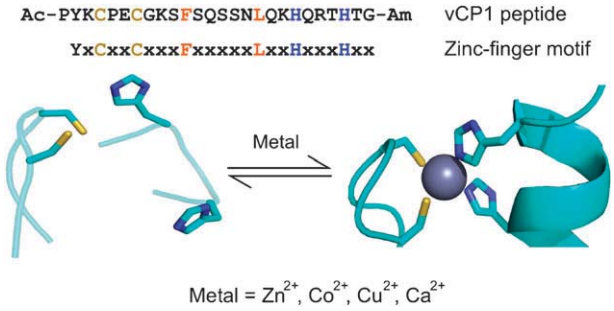


Fig. 1 The vCP1 peptide. Sequence, design criteria and proposed equilibrium in the presence of divalent metal ions. The sequences show key cysteine and histidine residues in yellow and blue, respectively. The residues shown in orange form a small hydrophobic core. The schematic shows the proposed rearrangement of a zinc finger when the two cysteine (sulfhydryl groups in yellow) and two histidine residues (imidazole nitrogen atoms in blue) coordinate a divalent metal ion (grey).[1] Adapted from PDB entries 1Z60 and 2DRP.

where K_0 is the reduced mobility (the measured mobility corrected to 273.15 K and 760 Torr), z is the ion charge state, e is the elementary charge, N is the gas number density, μ is the reduced mass of the ion-neutral pair, k_B is the Boltzmann constant and T is the gas temperature.

Here we use MS and IM-MS to examine a synthetic peptide for a consensus zinc finger sequence, vCP1, derived from the work of Berg *et al.*[7] Zinc fingers are found widely in Nature where they recognise DNA and perform other functions. The key zinc-binding residues are conserved cysteines (Cys, thiol containing) and histidines (His, imidazole containing), which coordinate metal and drive the folding of the polypeptide chain,[8] Fig. 1. vCP1 comprises just 26 amino acids making it a good model for studying peptide–metal interactions in solution and the gas phase.

To confirm the secondary structure and metal binding of vCP1 in solution, we used circular dichroism (CD) spectroscopy and ultraviolet (UV) spectroscopy, respectively. The CD spectra in the absence of metal ions confirm that the peptide was largely unstructured, whereas upon the addition of zinc the peptide folded to give a spectrum consistent with that reported for naturally occurring zinc fingers, Fig. 2a.

Addition of cobalt effected a similar change to that seen with zinc; in contrast, addition of copper and calcium changed the conformation of the peptide in a less pronounced manner (Supporting Information). The addition of cobalt to vCP1 was also followed by UV spectroscopy, Fig. 2b. The spectra are consistent with Cys_2His_2 tetrahedral coordination of the

[a] *School of Chemistry, University of Edinburgh, Edinburgh, UK. E-mail: Perdita.Barran@ed.ac.uk; Fax: 13 1650 7533; Tel: 13 1650 7533*
[b] *School of Chemistry, University of Bristol, Bristol, UK*
[c] *Department of Biochemistry, University of Bristol, Bristol, UK. E-mail: D.N.Woolfson@bristol.ac.uk; Fax: 11 7929 8611; Tel: 11 7954 6347*

† This article is part of the 'Emerging Investigators' themed issue for ChemComm.

‡ Electronic supplementary information (ESI) available: Experimental details for solution phase binding experiments; ion mobility measurements; MD calculations. See DOI: 10.1039/c0cc02445g

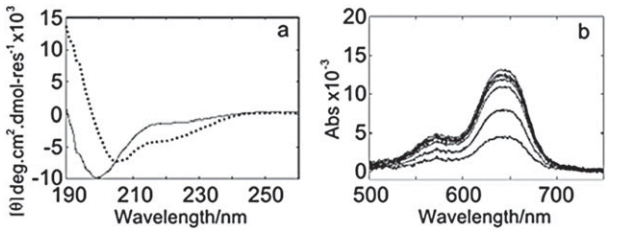


Fig. 2 Solution-phase spectra for vCP1. (a) CD spectra of 25 μM peptide in the absence (solid line) and presence (dotted line) of 50 μM zinc. (b) UV spectra for a titration of 4.4–47.4 μM Co^{2+} into 11.9 μM vCP1. Buffer: 5% isopropanol; pH 7.2; 20 mM ammonium acetate; 500 μM TCEP.

metal,[9] and a binding constant of 0.8 μM (see Supporting Information). A competitive-binding assay for zinc against the cobalt : vCP1 complex revealed sub-nM binding of zinc. These data confirm that vCP1 is a good model for metal binding studies under the solution conditions needed to best preserve intact complexes into the solvent-free environment of a mass spectrometer.

Fig. 3 shows typical nESI mass spectra for vCP1 without (Fig. 3a) and with (Fig. 3b–e) metals added; the regions for the +3 charge state(s) of the peptide are shown. The $[M + 3H]^{3+}$ or equally charged $[M + X + H]^{3+}$, where X = metal, were the dominant peaks in the spectra, ~8 times more intense than the +2 and +4 species. Use of TCEP (tris(2-carboxyethyl)-phosphine) maintained approximately 99% of the cysteines in the reduced state as evidenced by the isotopic cluster analysis. Theoretical fitting for the elemental compositions of the fully reduced peptide is superimposed on the experimental isotopic distribution, which allowed us to conclude that under these conditions the metals are coordinated by thiolates (S^-). Impurity peaks annotated on Fig. 3a are the common adducts of oxygen (*), sodium (**) and calcium (***). The ratio between the intensity of the ^{12}C peak for both apo and holo species was determined for each spectrum, which gave the order of metal affinity for apo-vCP1, $Zn^{2+} > Co^{2+} \gg Cu^{2+} \gg Ca^{2+}$, consistent with the solution-phase data.

Preferences for 1 : 1 vCP1 : metal binding were observed for Zn^{2+} (Fig. 3b), and to a slightly lesser extent for Co^{2+} (Fig. 3c) compared to the other metal ions. For both of these metals the same set of adducts, and at similar relative intensities as observed for apo-vCP1, were present with the holo-forms. A small fraction of vCP1 bound metal is seen in a 1 : 2 ratio. This was most pronounced for Cu^{2+}. Perhaps not surprisingly, in the presence of Cu^{2+} there was a higher fraction of the oxidised form of apo-vCP1, as revealed by the isotopic cluster analysis (Fig. 3d, inset), despite the 50-fold excess of TCEP.

We found Ca^{2+} to bind vCP1 with lowest affinity in both the gas and solution phase. A small amount of calcium (a contaminant of lab-ware, reagents, and even deionised water) was found to bind the peptide even in the presence of Zn^{2+}, leading us to conclude that Ca^{2+} binds vCP1 non-specifically, and probably away from the Cys_2His_2 binding site. We observed a low μM affinity for Zn^{2+} from MS analysis (see Supporting Information), considerably lower than found in solution. This difference may be attributable to

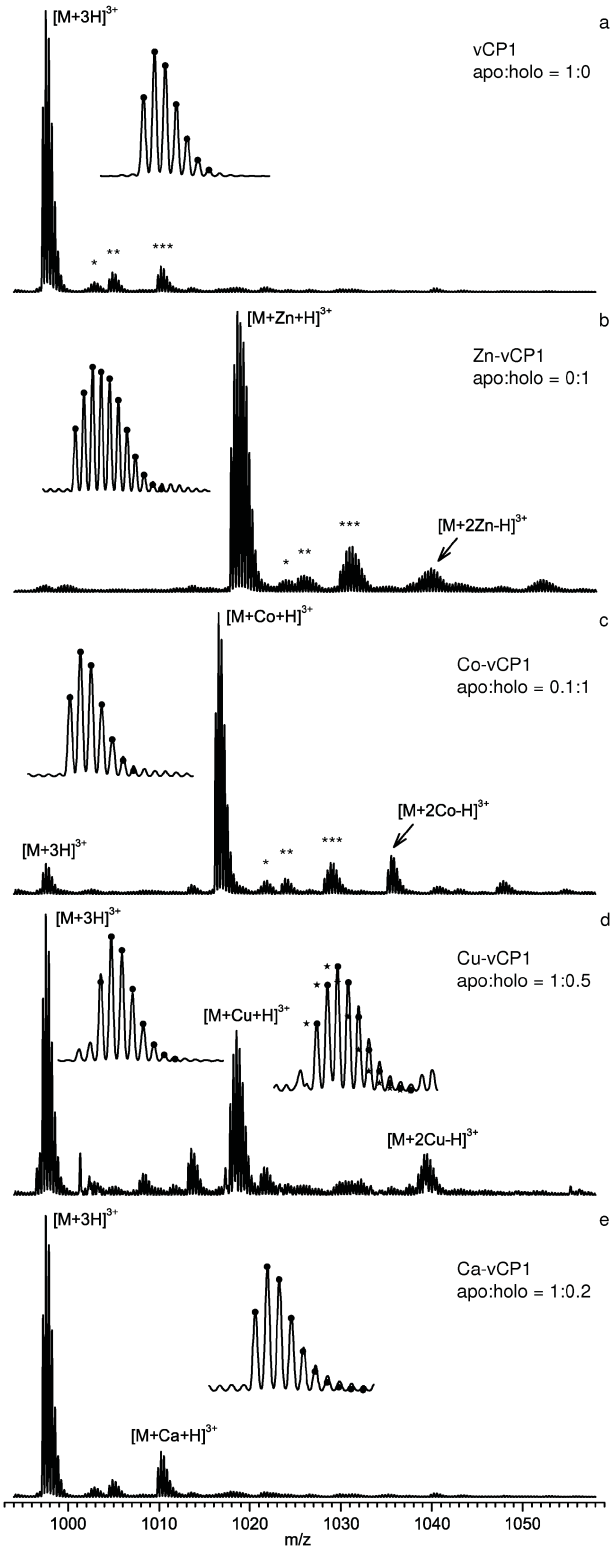


Fig. 3 nESI mass spectra of the +3 charge state of vCP1. For (a) the apo state, and with: (b) Zn^{2+}, (c) Co^{2+}, (d) Cu^{2+} and (e) Ca^{2+}. Ratios of apo to holo form were calculated from the intensities of the mono-isotopic peaks. Insets show the resolved isotopic clusters for the free and bound states along with theoretical fitting (●). Isotopic cluster analysis indicated that 99% of the peptide is reduced, except in the case of Cu^{2+}. Conditions: 20 μM peptide; 10 mM ammonium acetate; 5% isopropanol; pH 6.8; 200 μM TCEP; 100 μM metal acetate salts.

Table 1 Experimental collision cross sections of the +3 charge state of the apo and holo-vCP1 in 'buffered' conditions. The standard deviation for each measurement is given

	Apo-vCP1 (Å^2)	Holo-vCP1 (Å^2)
vCP1	488 ± 22	—
Zn-vCP1	487 ± 24	488 ± 21
Co-vCP1	518 ± 2	492 ± 17
Cu-vCP1	499 ± 12	493 ± 13
Ca-vCP1	528 ± 8	517 ± 2

desolvation effects, although the relative affinities for different metals are expected to be similar; this effect will be explored elsewhere.

By performing IM-MS experiments for all of the apo and the holo forms of vCP1 shown in Fig. 3, we obtain arrival time distributions (ATDs) at a range of drift voltages which are converted to collision cross sections following eqn (1) (Table 1). The packing found for the apo-vCP1 and Zn-vCP1 gave the lowest, and surprisingly similar, values. However all the values for the apo- and metal-bound forms were similar, and within experimental error. We expect to see more difference in collision cross section at elevated temperatures in the drift cell and thus obtain more information about the binding energies of this system in our future work.

To investigate these phenomena further, molecular dynamics (MD) simulations of vCP1 were performed. Fig. 4 shows representative nESI mass spectra and ATD data for the +3 charge state of Zn-vCP1, plus structures obtained by MD simulations.[10] The experimental collision cross sections of the apo-vCP1 and Zn-vCP1 compare favourably with those obtained from the MD giving confidence in both datasets. Interestingly, the apo form appears to possess some helical character (Fig. 4), although we find these helical regions are transient, particularly when compared to the holo form. When solvent is present in the simulations, however, there is little evidence for any helicity in the apo peptide (Supplementary Information).

In summary, we have illustrated how solution and gas-phase measurements can be combined to provide information on the binding stoichiometry and specificity of a model metal-binding peptide system. Qualitatively the agreement between the methods is excellent, and some semi-quantitative understanding can be gained also. Moreover, mass spectrometry studies have been extended to give collision cross sections, which give

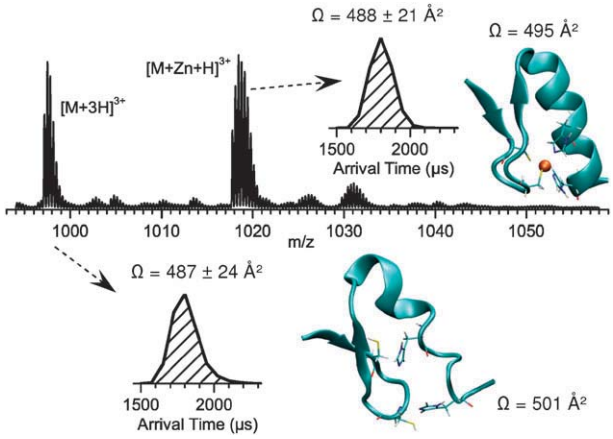


Fig. 4 nESI mass spectra of the +3 charge state of the 1 : 1 Zn : vCP1 species. Conditions are as for Fig. 3. Insets show the ATDs of apo and holo-vCP1, and representative MD structures, along with collision cross sections obtained experimentally and from the MD simulations.

further insight into coordination behaviour as well as the conformational preferences of a zinc-finger peptide. These results corroborate findings of MD studies on the peptide.

We acknowledge the support of the EPSRC, the RSC Analytical Division (who fund YB) and the HCP-Europa2 Scheme. CTA and ALB are funded by the BBSRC. We thank the Mann group (Bristol) for use of their spectrophotometer.

Notes and references

1 J. M. Berg, *Annu. Rev. Biophys. Biophys. Chem.*, 1990, **19**, 405–421.
2 T. Wyttenbach and M. T. Bowers, *Annu. Rev. Phys. Chem.*, 2007, **58**, 511–533.
3 B. J. McCullough, J. Kalapothakis, H. Eastwood, P. Kemper, D. MacMillan, K. Taylor, J. Dorin and P. E. Barran, *Anal. Chem.*, 2008, **80**, 6336–6344.
4 S. J. Valentine, A. E. Counterman and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 1999, **10**, 1188–1211.
5 M. De Cecco, E. S. Seo, D. J. Clarke, B. J. McCullough, K. Taylor, D. Macmillan, J. R. Dorin, D. J. Campopiano and P. E. Barran, *J. Phys. Chem. B*, 2010, **114**, 2312–2318.
6 T. Wyttenbach and M. T. Bowers, in *Modern Mass Spectrometry*, Springer, Berlin/Heidelberg, 2003, pp. 207–232.
7 C. A. Kim and J. M. Berg, *Nat. Struct. Biol.*, 1996, **3**, 940–945.
8 G. Parraga, S. J. Horvath, A. Eisen, W. E. Taylor, L. Hood, T. Y. Elton and R. E. Klevit, *Science*, 1988, **241**, 1489–1492.
9 Y. Shi, R. D. Beger and J. M. Berg, *Biophys. J.*, 1993, **64**, 749–753.
10 A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *J. Chem. Phys.*, 1998, **108**, 2416–2423.

